

Review

# AI and Wearables for Early Detection of Cognitive Impairment and Dementia: Systematic Review

Ander Cejudo<sup>1,2</sup>, MSc; Markel Arrojo<sup>1</sup>, MSc; Cristina Martín<sup>1,2,3</sup>, PhD; Aitor Almeida<sup>2</sup>, PhD

<sup>1</sup>Vicomtech Foundation, Basque Research and Technology Alliance (BRTA), Donostia-San Sebastián, Spain

<sup>2</sup>Faculty of Engineering, University of Deusto, Bilbao, Spain

<sup>3</sup>Biodonostia Health Research Institute, e-Health Group, San Sebastián, Spain

**Corresponding Author:**

Ander Cejudo, MSc

Vicomtech Foundation, Basque Research and Technology Alliance (BRTA)

Mikeletegi 57

Donostia-San Sebastián, 20009

Spain

Phone: 34 943 309 230

Email: [acejudo@vicomtech.org](mailto:acejudo@vicomtech.org)

## Abstract

**Background:** Traditional cognitive screening relies on episodic clinical assessments and may miss early changes preceding cognitive impairment and dementia. Wearable and mobile health technologies enable continuous monitoring of sleep, physical activity, and circadian rhythms, generating digital biomarkers that may support scalable early detection and prevention. However, current evidence remains fragmented across devices, analytic approaches, and cognitive outcomes.

**Objective:** This study synthesizes and critically evaluates recent evidence on wearable devices for early detection and prevention of cognitive impairment and dementia, focusing on device categories, cognitive outcomes, analytic approaches, and prevention relevance.

**Methods:** We searched PubMed, Scopus, ACM Digital Library, and SpringerLink for peer-reviewed studies published between January 2020 and December 1, 2025. Eligible studies included human participants with a mean age  $\geq 50$  years, continuous wearable-derived data collected for  $\geq 24$  hours, and validated cognitive outcomes; reviews, protocols, smartphone-only studies, and pharmacological interventions were excluded. Two reviewers independently screened studies, extracted data, and assessed risk of bias using the Appraisal Tool for Cross-Sectional Studies, Newcastle-Ottawa Scale, Cochrane Risk of Bias tool, and Quality Assessment of Diagnostic Accuracy Studies-2. Owing to substantial heterogeneity in devices, outcomes, and analytic methods, quantitative meta-analysis was not feasible; a structured narrative synthesis was conducted in accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidance. This study was not prospectively registered.

**Results:** We included 49 studies, with sample sizes ranging from 14 to 91,948 participants ( $>200,000$  total) and a median sample size of 145. Most used research-grade actigraphy (43/49, 87.8%), while fewer used commercial wearables (7/49, 14.3%). Cognitive outcomes most frequently relied on global screening instruments, including the Mini-Mental State Examination (18/49, 36.7%), followed by *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*)-based clinical diagnoses (7/49, 14.3%) and the Montreal Cognitive Assessment (7/49, 14.3%). Analytic approaches were predominantly statistical (36/49, 73.5%), with fewer studies applying machine learning (7/49, 14.3%) or deep learning methods (6/49, 12.2%). Statistical analyses linked disrupted sleep, circadian rhythm fragmentation, and irregular activity patterns to worse cognitive outcomes, with modest-to-moderate effect sizes. Machine learning and deep learning approaches reported classification performance with area under the curve values between approximately 0.70 and 0.95. Approximately one-quarter of the studies (13/49, 26.5%) addressed early detection or prevention through longitudinal risk estimation or predictive modeling. Key limitations included small sample sizes, short monitoring durations, and limited external validation.

**Conclusions:** Wearable-derived behavioral markers show promise for early risk stratification. This review advances the field by shifting from descriptive associations toward a digital phenotyping framework evaluating artificial intelligence-driven prediction in the preclinical window. Unlike prior reviews focused on established dementia, it differentiates direct predictive evidence from

indirect correlational findings and critically assesses methodological maturity. Continuous, passive monitoring may enable scalable detection of subtle behavioral changes, supporting earlier and more personalized risk reduction strategies.

(*J Med Internet Res* 2026;28:e86262) doi: [10.2196/86262](https://doi.org/10.2196/86262)

## KEYWORDS

actigraphy; artificial intelligence; circadian rhythms; cognitive impairment; dementia; machine learning; mHealth; mobile health; prevention; sleep; wearable devices

## Introduction

### Background

Cognitive decline and dementia represent a major public health concern, with significant global and regional prevalence. In 2021, approximately 57 million people were affected by dementia worldwide, with more than 60% residing in low- and middle-income countries [1]. As the global population ages, dementia is expected to impact 78 million individuals by 2030 and 139 million by 2050 [2]. This increasing prevalence is attributed to both the rising number of older adults and the fact that dementia disproportionately affects the older population. Furthermore, the severity of the problem is anticipated to rise more sharply in regions with significant demographic changes, particularly low- and middle-income countries, where health care resources are often limited [3].

Dementia severely affects the quality of life, as it is a leading cause of disability and dependency in older individuals [4]. In addition to the physical and cognitive impairments experienced by individuals, dementia imposes a heavy emotional and economic burden on families and caregivers. Caregivers, often family members, are responsible for a substantial amount of care, with an average of 5 hours of daily care and supervision per individual affected. The global cost of dementia is estimated to exceed US \$1.3 trillion, with informal caregiving accounting for approximately half of this expenditure. Health care and social systems face immense strain in responding to the needs of individuals with dementia, highlighting the urgent need for preventive measures and early intervention [2].

There are several modifiable risk factors for cognitive decline and dementia, including physical inactivity, unhealthy diets, smoking, and cardiovascular conditions such as hypertension, diabetes, and obesity [5]. Early identification of cognitive decline is crucial for enabling timely interventions. Traditional screening methods, such as the Mini-Mental State Examination (MMSE) and Montreal Cognitive Assessment (MoCA) [6], have limitations due to their episodic nature and reliance on periodic visits to health care providers. These methods are often influenced by situational factors [7,8] and may not detect early signs of decline or track changes in cognitive function over time [9,10]. In addition, performance on these instruments can be affected by educational level, language, and cultural background, and they provide only a cross-sectional snapshot of cognitive status [11]. Therefore, there is a growing need for continuous and more sensitive methods of monitoring cognitive health [12]. At the same time, the use of artificial intelligence (AI)-driven wearable technologies for continuous cognitive monitoring raises ethical and data privacy considerations, as such systems may involve the collection of sensitive behavioral and

health-related data, underscoring the importance of appropriate data governance and responsible implementation [13].

Wearable devices present significant potential for the continuous and passive monitoring of key physiological variables, including activity levels, sleep patterns, heart rate, and gait. These devices can detect subtle, longitudinal changes in behavior that may occur before clinical symptoms emerge [14], allowing for the early identification of cognitive decline. AI enhances this process by analyzing data from wearables to identify patterns that indicate potential risks. AI algorithms can process large volumes of data, detecting even small shifts in physiological and behavioral patterns, which can inform personalized interventions. By leveraging AI for predictive analytics, it is possible to improve early detection capabilities, facilitating timely interventions and supporting more effective management of cognitive health [15]. This combination of wearables and AI offers a promising approach [16] to enhancing dementia prevention and care.

### Objective

Previous reviews have mainly examined associations between wearable-derived measures and dementia outcomes [16], often focusing on physiological or behavioral differences in individuals already diagnosed with the disease. Other reviews have emphasized specific domains, such as gait and mobility assessment using wearables [17] or the relationship between physical activity and cognitive decline [18]. In contrast, this review focuses on how wearable-derived data are analyzed, with particular attention to advanced analytic approaches, including machine learning and deep learning, used to support early detection and prevention across the cognitive aging spectrum, from subjective cognitive decline to clinical dementia. By synthesizing evidence within a digital phenotyping framework, this review highlights how longitudinal, multimodal wearable data combined with contemporary analytic methods can move beyond descriptive associations toward individualized risk characterization and prevention-oriented applications.

Thus, the main objective of this systematic review is to synthesize and critically evaluate the current evidence base and methodological maturity of wearable devices for the early detection and prevention of cognitive impairment and dementia. We structure our synthesis around 4 specific objectives:

- To evaluate the usage trends and application contexts of wearable technologies, specifically comparing the deployment of research-grade actigraphy vs consumer-grade devices.
- To synthesize the strength of statistical associations between wearable-derived digital biomarkers (eg, sleep and circadian rhythms) and standard clinical cognitive assessments.

- To critically assess the performance and robustness of analytic approaches, comparing the outcomes of conventional statistical modeling, machine learning, and deep learning techniques.
- To assess how early detection is addressed in wearable-based studies by differentiating between studies that directly implement early detection and those that frame their findings as potentially relevant to early detection.

By addressing these specific objectives, the review aims to clarify the current state of knowledge and identify priorities for advancing wearable-based approaches toward clinical and public health application.

## Methods

### Overview

The reporting of this systematic review was guided by the standards of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement [19] (Multimedia Appendix 1). The literature search strategy was developed and reported in accordance with the PRISMA-S (Preferred Reporting Items for Systematic Reviews and Meta-Analyses literature search extension) [20] to enhance transparency and reproducibility of the search process (Multimedia Appendix 2). The aim was to evaluate the association between wearable device-based sleep and activity measures and the onset or progression of cognitive impairment, including mild cognitive impairment (MCI), Alzheimer disease, and cognitive decline. The review process was carried out in multiple stages: identification of relevant articles, screening for eligibility, data extraction, and assessment of risk of bias. Studies that meet predefined inclusion criteria were included, and their findings were synthesized to provide insights into the role of wearable devices in monitoring cognitive health. No review protocol was prepared or registered before conducting this study.

### Information Sources and Search Strategy

A comprehensive search strategy was used across multiple electronic databases, including PubMed, Scopus, ACM Digital Library, and SpringerLink, to identify relevant studies published after 2020. The search was conducted using a combination of MeSH (Medical Subject Headings) terms and keywords related to cognitive impairment, sleep and activity measures, wearable devices, and their associations with cognitive outcomes. Databases were searched individually using their native interfaces. Duplicate records were removed, and studies published before 2020 were discarded to align the scope of this systematic review with contemporary research questions, analytic frameworks, and the use of wearable devices

representative of current practice in digital phenotyping and early detection and prevention of cognitive decline. The search strategy was designed to capture a broad range of studies, including randomized controlled trials, cohort studies, and observational studies, to provide a comprehensive overview of the existing evidence.

The search strategy integrates key concepts relevant to studies on the early detection of cognitive impairment and dementia using wearable devices that monitor activity and sleep. Table 1 lists the search terms, organized into categories capturing cognitive impairment, sleep and activity measurements, wearable devices, and their application in early detection. Within each group, similar terms were combined using the OR operator, and the 4 groups were combined using the AND operator. Conceptually, the search followed the structure: (HC<sub>1</sub> OR HC<sub>2</sub> OR ...) AND (Data<sub>1</sub> OR Data<sub>2</sub> OR ...) AND (Wearable<sub>1</sub> OR Wearable<sub>2</sub> OR ...) AND (EarlyDetection<sub>1</sub> OR EarlyDetection<sub>2</sub> OR ...). Full search queries are included in Multimedia Appendix 3.

Alzheimer disease is specifically included under dementia, as it accounts for 60%-70% of all dementia cases [1]. Data capture must include information on sleep, activity, or physiological signals. Wearable devices may be commercial (eg, smartwatches) or noncommercial (ad hoc or research-grade devices) and are typically worn on the body for extended periods. This review focuses on studies advancing early detection and prevention, excluding research that examines features of the condition without the potential to identify cognitive decline before clinical diagnosis or progression to a more advanced stage.

To consider an article in this study, at least 1 search term of each group must be present in either the title or the abstract. This ensures that the manuscript has a clear focus on continuous data collection on sleep, activity, or physiological signals and links these measurements to cognitive impairment outcomes. The search was limited to studies published after 2020 and in English to ensure accessibility to full-text articles.

No additional search methods were used, and no study registries were searched. No targeted website searching or manual browsing of conference proceedings was conducted. Study authors and experts were not contacted for additional data. No published search filters were used. Search strategies were developed specifically for this review and were not adapted from prior reviews. No search updates were performed after the initial search, and the search strategy was not peer reviewed. Reference lists of included studies were screened to identify additional relevant records.

**Table 1.** Groups of search terms and representative keywords used to identify studies in the systematic review, organized by health condition, data type, wearable technology, and early detection.

Group	Search terms
Health condition	Cognitive impairment, mild cognitive impairment, dementia, cognitive decline, Alzheimer, neurocognitive disorder, and memory impairment
Data	Sleep, sleep duration, sleep quality, circadian rhythm, rest-activity rhythm, physical activity, actigraphy, accelerometry, heart rate, heart rate variability, HRV, respiratory rate, body temperature, respiration, skin temperature, step count, steps, and distance
Wearable	Wearable, wearable device, wearable devices, wearable sensor, wearable sensors, body-worn sensor, body-worn sensors, fitness tracker, smartwatch, smart band, wristband, wristbands, accelerometers, consumer-grade wearable, actigraph, actigraphy, ambulatory monitor, and digital biomarker
Early detection	Risk factor, onset, association, predictor, correlation, early detection, screening, and prevention

## Eligibility Criteria

The inclusion and exclusion criteria that guided the selection of studies for this systematic review are presented in [Textbox 1](#). The criteria are designed to ensure that the review captures only the most relevant, recent, and methodologically sound research.

For inclusion, studies must be published between 2020 and 2025 to reflect the most recent developments in the field. Only publications identified by the predefined search queries were considered. Eligible studies must present either new statistical outcomes or AI-based results that provide original contributions. Because the review focuses on prevention and early detection, manuscripts must directly address these topics. Furthermore, included studies were required to involve human participants with a mean age of 50 years or older, ensuring relevance to an aging population. Another important condition was the continuous collection of data from wearable devices for at least

24 hours, which allows for robust and objective measurements. Only peer-reviewed journal articles were considered, discarding gray literature and preprints.

Exclusion criteria further refine the selection. Publications not written in English or not available in full text, for example, when behind an inaccessible paywall, were excluded. Review articles, study protocols, and books were not considered to maintain a focus on original research. Retracted publications were excluded. In addition, studies were not eligible if they did not apply validated measures of cognitive impairment or dementia, focused on pharmacological interventions, relied exclusively on self-reported data or smartphone-based sensing, or included only healthy participants without relevant clinical characteristics. No minimum sample-size threshold was applied during study selection. As a result, included studies span a wide range of sample sizes, including small-sample investigations (eg,  $n < 30$ ).

**Textbox 1.** Inclusion and exclusion criteria applied in the systematic review.

### Inclusion criteria:

- Published between 2020 and 2025
- Identified by the search queries
- Articles presenting new statistical outcomes or artificial intelligence results
- Manuscripts closely related to prevention and early detection
- Human participants must be involved in the study with average age of 50 years or older
- Continuous data captured from wearable devices for 24 hours or more

### Exclusion criteria:

- Publication not in English
- Publication behind a paywall that cannot be retrieved
- Review papers, protocols, and books
- Retracted publications
- Validated measures of cognitive impairment and dementia not included
- Focus on drug or pharmacological interventions
- Only used self-reported data or smartphone-based sensing
- Only included healthy participants

## Risk of Bias Assessment

The risk of bias in included studies was assessed using the Appraisal Tool for Cross-Sectional Studies, the Newcastle-Ottawa Scale for cohort studies, the Cochrane Risk of Bias tool for randomized controlled trials, and the Quality Assessment of Diagnostic Accuracy Studies-2 for diagnostic studies [21]. These instruments evaluate domains such as selection, performance, detection, attrition, and reporting bias, as well as other potential sources of bias. Each study was independently evaluated by 2 coauthors (AC and MA), with disagreements resolved through discussion or consultation with a third coauthor (CM). The overall risk of bias was classified as low, moderate, or high, and these assessments were considered when interpreting the results of the review.

## Study Selection and Data Collection Process

Search results from each database were exported as BibTeX files and managed in Zotero (Corporation for Digital Scholarship), then converted to comma-separated values for automated preprocessing using a custom Python script. This process harmonized metadata across databases, removed duplicates, applied publication year filters, and verified the presence of predefined search terms. Automated preprocessing was conducted by the first author and independently verified by the second author. The following data items were extracted from each eligible study: bibliographic details (author and year), study design, participant characteristics (sample size, age, and diagnosis), wearable device specifications (type, brand, and sensor modality), monitoring duration, cognitive assessment tools, analytic methods (statistical vs AI), and primary outcomes (statistical associations or predictive performance metrics). Study selection followed a 3-stage screening process, with independent title, abstract, and full-text reviews performed by the first and second authors. Discrepancies were resolved through consensus or adjudication by a third author. Eligible studies were synthesized using a structured narrative approach aligned with predefined review objectives, grouping studies by device type, cognitive outcomes, analytic methods, and prevention-oriented contributions.

Analytic approaches were classified according to the overall modeling paradigm rather than individual algorithms. Studies were categorized as AI-based when wearable-derived data were analyzed within predictive modeling frameworks emphasizing automated pattern learning, such as machine learning or deep learning methods, including regression-based models implemented within machine learning pipelines (eg, automated

feature selection or cross-validation). In contrast, studies relying on predefined statistical models for inferential or association analyses were classified as statistical model-based approaches. In this review, “direct” evidence was defined for studies demonstrating individual-level early detection capability, including both cross-sectional screening or discriminative analyses and longitudinal or interventional designs enabling risk estimation or prevention. Cross-sectional studies were therefore considered direct when they demonstrated the ability to distinguish preclinical or prodromal disease states at the time of assessment.

## Synthesis Methods

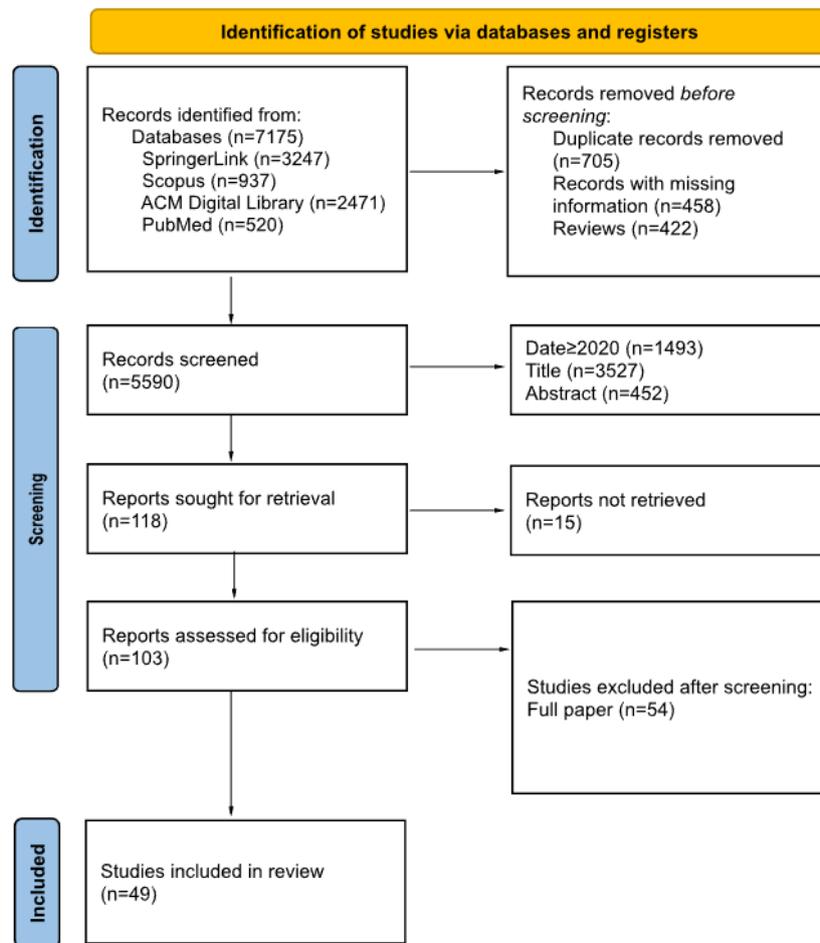
Quantitative synthesis was planned only for outcomes estimating a common construct using comparable analytical objectives and metrics. When outcomes cannot be mapped to a shared estimand for early detection, quantitative pooling was not conducted. In such cases, results were synthesized using a structured narrative approach aligned with the predefined review objectives. This approach grouped studies according to analytic strategy, cognitive assessment, and wearable measurement characteristics, without applying formal qualitative synthesis methodologies.

## Results

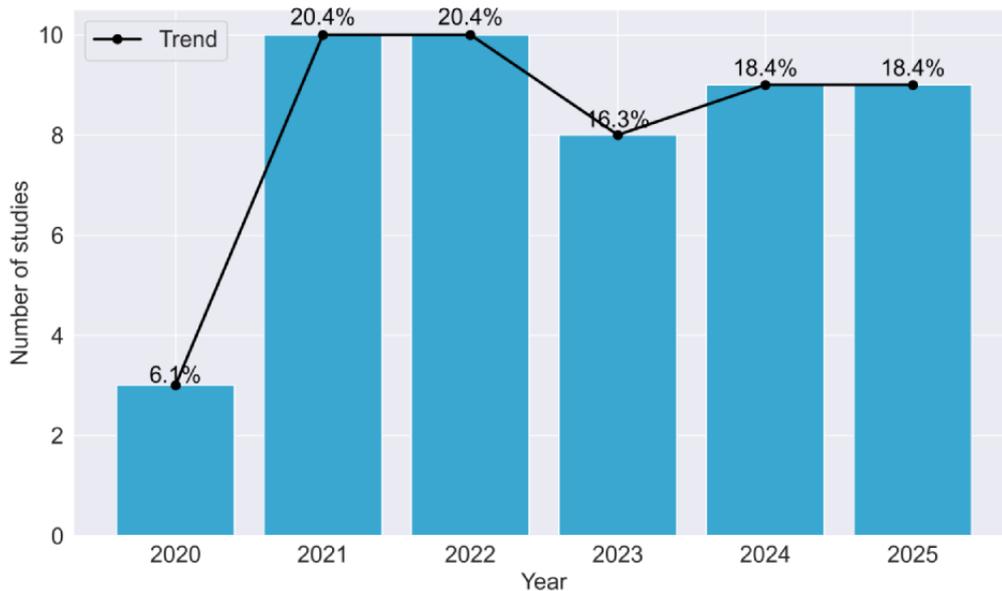
### Study Selection

The database search identified 7175 records (SpringerLink, n=3247; Scopus, n=937; ACM Digital Library, n=2471; and PubMed, n=520). After removal of 705 duplicates, 458 records with missing information, and 422 reviews, a total of 5590 records remained for screening. Of these, 1493 were excluded based on publication year (<2020, n=1493), title (n=3527), or abstract (n=452). The remaining 118 reports were sought for full-text retrieval, but 15 could not be obtained. A total of 103 full-text articles were assessed for eligibility, and 54 were excluded for not meeting the inclusion criteria. Finally, a total of 49 studies were included in this systematic review (Figure 1). Figure 2 shows the number of included studies by publication year. At the full-text screening stage, most studies were excluded because they did not include continuous wearable-based data collected over at least 24 hours, did not use wearable-derived data as part of the analysis, or did not report results relevant to the early detection of cognitive impairment or dementia. Table 2 shows the relevant information extracted from the included studies, and additional details are provided in Multimedia Appendix 4.

**Figure 1.** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the study selection process.



**Figure 2.** Temporal distribution of included studies by publication year (2020-2025). The bar chart and overlaid trend line illustrate the progression of research interest in the field. Data show a substantial increase in publication volume beginning in 2021, with research activity remaining consistently elevated through the 2024-2025 period. Data for 2025 are incomplete by 1 month.



**Table 2.** Summary of included cross-sectional studies investigating wearable-derived measures in relation to dementia outcomes. The table reports study source, cross-sectional study design, participant groups and mean ages, and assessed risk of bias.

Source	Objective	Groups, n	Age (years), mean (SD)	Risk of bias
Lentzen et al [22]	Early detection	HC <sup>a</sup> =69; DEM <sup>b</sup> =160	69.2 (7.7)	Low
Jones et al [23]	Early detection	NCD <sup>c</sup> =21; MCI <sup>d</sup> =19	67.2 (8.4)	Low
Satomi et al [24]	Early detection	HC=25; MCI=63; DEM=21	79.3 (6.3)	Low
Sauers et al [25]	Risk assessment	HC=196; MCI=47	74	Low
Baril et al [26]	Risk assessment	HC=159; MCI=44	68.25	Low
Seo et al [27]	Early detection	HC=1519; MCI=353; DEM=87	70.1 (6.1)	Low
Huang et al [28]	Preventive intervention	HC=225; MCI=133; DEM=36	71.3(7.8); 72.9 (8.0)	Low
Palmer et al [29]	Progression monitoring	SCI <sup>e</sup> =29; MCI=79	68.8 (8.9)	Low
Lim [30]	Early detection	HC=14; DEM=4	≥65	Low
Boa Sorte Silva et al [31]	Early detection	MCI=113	73.1 (5.7)	Low
Kim et al [32]	Progression monitoring	HC=41; NCD=54	Median 60	Low
Basta et al [33]	General insight	MCI=40; DEM=49	78.3 (6.5)	Low
Roh et al [34]	Progression monitoring	MCI=70; DEM=30	Median 73	Low
Liu et al [35]	Early detection	HC=31; MCI=68	66.5	Low
Espinosa et al [36]	Early detection	SCI=44; MCI=99	67.4 (7.9)	Medium
Manning et al [37]	Risk assessment	HC=7; MCI=2; DEM=5	65.8	Medium
Ghosal et al [38]	Early detection	HC=54; DEM=38	73.4 (7.1)	Medium
Park et al [39]	Early detection	MCI/DEM=177; HC=14,482	74.1 (7); 39 (2.3)	Medium
Palmer et al [40]	Early detection	SCI=23; MCI=44	65.8 (7.9)	Medium
Wei et al [41]	Early detection	HC=20; DEM=8	Mean 59.2	Medium
Corbi and Burgos [42]	Early detection	HC=37; DEM=74	80.5	Medium

<sup>a</sup>HC: healthy controls.

<sup>b</sup>DEM: dementia.

<sup>c</sup>NCD: neurocognitive disorder.

<sup>d</sup>MCI: mild cognitive impairment.

<sup>e</sup>SCI: subjective cognitive impairment.

### Characteristics of Included Studies

A total of 49 studies were included in the review (summarized in [Tables 3-8](#)). Sample sizes ranged from 14 to 91,948 participants, with a median of 145. A total of 5 large cohorts heavily influenced the mean values (ranging between 47,371 and 91,948 participants). When stratified by analytic approach, studies using traditional statistical methods (36/49, 73.5%) had a median of 100 participants (mean 2597, SD 13,574), while those applying machine learning (7/49, 14.3%) had a median of 99 participants (mean 114.9, SD 76.4). Studies using deep learning (6/49, 12.2%) had a median of 108 participants (mean 397.6, SD 465.4), largely driven by population-scale datasets. Most studies were rated as low risk (34/49, 69.4%), nearly one-third (13/49, 26.5%) were classified as medium risk, and

2/49 (4.1%) as high risk. Participant populations were predominantly composed of individuals with MCI (29/49, 59.2%), dementia (35/49, 71.4%), or healthy controls (36/49, 73.5%). A smaller subset included participants with subjective cognitive decline (4/49, 8.2%) or other at-risk groups (2/49, 4.1%). Regarding study design, half of the studies were cross-sectional (21/49, 42.9%), followed by longitudinal cohorts (24/49, 49%), with fewer randomized controlled trials or diagnostic studies. Most studies (28/49, 57.1%) focused on early detection, while others assessed risk factors (8/49, 19.5%), monitored disease progression (7/49, 14.3%), or preventive interventions (5/49, 10.2%). Risk of bias (complete tables are provided in [Multimedia Appendix 5](#)) was generally low (35/49, 71.4%), though higher risk ratings were associated with small sample sizes and lack of external validation.

**Table 3.** Summary of included cohort studies investigating wearable-derived measures in relation to dementia outcomes. The table reports study source, study objective, cohort study design, participant groups and mean ages, and assessed risk of bias.

Source	Objective	Groups, n	Age (years), mean (SD)	Risk of bias
Baril et al [43]	Progression monitoring	HC <sup>a</sup> =159; MCI <sup>b</sup> =44	68.2 (5.4)	Low
Sun et al [44]	Early detection	HC=807; DEM <sup>c</sup> =270	80.9 (7.3)	Low
Winer et al [45]	Early detection	HC=82,377; DEM=452	62.0 (7.8)	Low
Haghighyegh et al [46]	Early detection	HC=90,962; MCI or DEM=555	62.4 (7.8)	Low
Basta et al [47]	Progression monitoring	HC=29; MCI=49; DEM=32	80.3 (6.6)	Low
Milton et al [48]	Risk assessment	HC=476; MCI=164; DEM=93	82.5 (2.9)	Low
Skourti et al [49]	Progression monitoring	HC=146; MCI=231; DEM=128	72.8 (6.7)	Low
Gao et al [50]	Early detection	HC=409; MCI=61; DEM=81	58 (8)	Low
Jeon et al [51]	Early detection	HC=100; MCI=29	69.3 (7.7)	Low
Yi Lee et al [52]	Preventive intervention	HC=123; MCI=51	75.6 (6.9)	Low
Lysen et al [53]	General insight	HC=1262; DEM=60	66.1 (7.6)	Low
Agudelo et al [54]	Early detection	DEM=1035	55.2 (2.5)	Low
Ning et al [55]	Risk assessment	HC=91,212; DEM=736	Median 63	Low
Ning et al [56]	Risk assessment	HC=89,619; DEM=710	Median 64.9	Low
Zhao et al [57]	Risk assessment	HC=87,857; DEM=735	61.9 (7.9)	Low
Lu et al [58]	Early detection	HC=303; MCI=130; DEM=277	81.1 (5.2)	Low
Chan et al [59]	Early detection	HC=46,984; DEM=387	67.0 (4.0)	Low
Shi et al [60]	Early detection	HC=346; DEM=344	70.1 (6.9)	Low
Xiao et al [61]	Progression monitoring	HC=570; MCI=120; DEM=73	84.1	Low
Hoepel et al [62]	Preventive intervention	HC=1849; DEM=50	71.3 (9.26)	Low
Cho et al [63]	Early detection	DEM=222	80.4 (7.4)	Medium
Plotogea et al [64]	Early detection, preventive intervention	HC=25; SCI <sup>d</sup> =7; MCI=17; DEM=25	58.9 (9.8)	Medium
Targa et al [65]	Early detection	DEM=100	Median 76	Medium
Cho et al [66]	Risk assessment	DEM=145	81.2 (6)	Medium

<sup>a</sup>HC: healthy controls.

<sup>b</sup>MCI: mild cognitive impairment.

<sup>c</sup>DEM: dementia.

<sup>d</sup>SCI: subjective cognitive impairment.

**Table 4.** Summary of included studies with other study designs investigating wearable-derived measures in relation to dementia outcomes. The table reports study source, study objective (early detection or progression monitoring), study design (diagnostic study, randomized controlled trial, or quasi-experimental study), participant groups and ages, and assessed risk of bias.

Source	Objective	Study design	Groups, n	Age (years)	Risk of bias
Rykov et al [67]	Early detection	Quasi-experimental study	MCI <sup>a</sup> =17	60.3 (SD 4.5)	Medium
David et al [68]	Progression monitoring	Randomized controlled trial	DEM <sup>b</sup> =38	70 (SD 7)	Medium
Khosroazad et al [69]	Early detection	Diagnostic study	HC=45; MCI <sup>c</sup> =50	73.6	High
Hossain et al [70]	Early detection	Diagnostic study	HC=22; MCI=6; DEM=5	57.5	High

<sup>a</sup>HC: healthy controls.

<sup>b</sup>DEM: dementia.

<sup>c</sup>MCI: mild cognitive impairment.

**Table 5.** Wearable devices, recording parameters, and analytic methods used across the included studies. The table lists device type, monitoring duration, product and model, analytic approach (statistical, machine learning, deep learning, or clustering), primary task, and cognitive assessments applied.

Source	Wearable category	Wearable	Product (company name)	Duration (days), n	Model	Cognitive test
Espinosa et al [36]	Research	Actigraph	Actiwatch Spectrum (Philips)	13.7	Statistical	MMSE <sup>a</sup>
Lentzen et al [22]	Commercial research	Smartwatch	Fitbit Charge 3 (Google); AX3 (Axivity); Physilog (Gait Up)	56	Machine learning	Amsterdam IADL <sup>b</sup>
Manning et al [37]	Research	Actigraph	Actiwatch (Philips)	14	Statistical	MoCA <sup>c</sup>
Jones et al [23]	Commercial	Smartwatch	Apple Watch 8 (Apple)	6.9	Statistical	HVLT <sup>d</sup>
Satomi et al [24]	Research	Actigraph	ActTrust (Condor Instruments)	7	Statistical	CDR <sup>e</sup>
Baril et al [43]	Research	Actigraph	Actiwatch (Philips)	6.5	Statistical	RBANS <sup>f</sup>
Sauers et al [25]	Research	Actigraph	Sleep Profiler (Advanced Brain Monitoring); Actiwatch2 (Philips); Alice PDx (Philips)	6	Statistical	CDR
Sun et al [44]	Research	Actigraph	Actical (Philips)	14	Deep learning	NINCDS-ADRDA <sup>g</sup>
Winer et al [45]	Research	Accelerometer	AX3 (Axivity)	7	Statistical	TMT <sup>h</sup>
[67]	Multimodal research	Wrist-worn device	E4 wristband (Empatica)	10	Machine learning	NTB <sup>i</sup>
Haghighayegh et al [46]	Research	Accelerometer	AX3 (Axivity)	6	Statistical	ICD-10 <sup>j</sup>
Basta et al [47]	Ad hoc	Actigraph	Not specified	7	Statistical	ICD-10
Baril et al [26]	Research	Actigraph	Actiwatch (Philips)	6.5	Statistical	MMSE
Milton et al [48]	Multimodal research	Actigraph	SleepWatch-O (Ambulatory Monitoring)	3	Statistical	MMSE
Cho et al [63]	Research	Actigraph	ActiGraph wGT3X-BT (ActiGraph)	14	Machine learning	MMSE
Seo et al [27]	Research	Accelerometer	HW-100 (Kao Corporation)	28	Statistical	MMSE
Huang et al [28]	Research	Actigraph	GENEActiv Original (Activinsights Company)	6	Statistical	MMSE
Khosroozad et al [69]	Research	Actigraph	Actiwatch (Philips)	7	Deep learning	HVLT
Skourti et al [49]	Research	Actigraph	Actigraph GT3XP (ActiGraph)	3	Statistical	RAVLT <sup>k</sup>
David et al [68]	Commercial	Smartwatch	Fitbit Charge 2 (Google)	13	Statistical	MoCA
Gao et al [50]	Research	Accelerometer	AX3 (Axivity)	7	Statistical	ICD-10
Ghosal et al [38]	Research	Accelerometer	Actigraph GT3x+ (ActiGraph)	7	Machine learning	CDR
Palmer et al [29]	Research	Actigraph	Actiwatch Spectrum (Philips)	10.5	Deep learning	MMSE
Jeon et al [51]	Research	Actigraph	Actiwatch 2 (Philips)	5.5	Statistical	MMSE
Lim et al [30]	Ad hoc	Smartwatch	Not specified	7	Deep learning	MMSE
Plotogea et al [64]	Research	Actigraph	Actiwatch Spectrum Pro (Philips)	7	Statistical	PHES <sup>l</sup>
Boa Sorte Silva et al [31]	Research	Actigraph	MotionWatch8 (CamNtech)	7	Statistical	MoCA
Park et al [39]	Commercial	Accelerometer	Fitmeter (FitNLife); ActiGraph AM-7164 (ActiGraph)	24.6 7	Deep learning	MMSE; CDR; GDS <sup>m</sup>
Palmer et al [40]	Research	Actigraph	Actiwatch Spectrum (Philips)	7	Statistical	MMSE
Kim et al [32]	Research	Actigraph	Spectrum Plus (Philips); Actiwatch Spectrum (Philips)	10.5	Statistical	NIH <sup>n</sup>

Source	Wearable category	Wearable	Product (company name)	Duration (days), n	Model	Cognitive test
Targa et al [65]	Research	Actigraph	Actiwatch 2 (Philips)	14	Statistical	MMSE
Wei et al [41]	Commercial	Smartwatch	Mi Band 2 (Xiaomi)	14	Statistical	MMSE
Hossain et al [70]	Ad hoc	Smartwatch	Not specified	182.5	Machine learning	MMSE
Cho et al [66]	Ad hoc	Actigraph	wGT3X-BT (ActiGraph)	11.5	Statistical	MMSE
Corbi and Burgos [42]	Research	Accelerometer	AX3 (Axivity)	4	Machine learning	GDS
Basta et al [33]	Research	Actigraph	Actigraph GT3XP (ActiGraph)	3	Statistical	MMSE
Yi Lee et al [52]	Research	Actigraph	GENEActiv Original (ActivInsights)	7	Statistical	MoCA
Roh et al [34]	Commercial	Accelerometer	Fitmeter (FitNLife)	4	Statistical	SNSB <sup>o</sup>
Lysen et al [53]	Research	Actigraph	ActiWatch AW4 (CamNtech)	6	Statistical	MMSE
Agudelo et al [54]	Research	Actigraph	Actiwatch Spectrum (Philips)	7	Statistical	B-SEVLT-Sum <sup>p</sup>
Ning et al [55]	Research	Accelerometer	AX3 (Axivity)	7	Statistical	ICD-10
Ning et al [56]	Research	Accelerometer	AX3 (Axivity)	7	Statistical	ICD-10
Zhao et al [57]	Research	Accelerometer	AX3 (Axivity)	7	Statistical	ICD-10
Lu et al [58]	Research	Actigraph	wGT3X-BT (Ametris)	7	Statistical	MoCA
Chan et al [59]	Research	Smartwatch	Not specified	7	Statistical	ICD-10
Shi et al [60]	Research	Actigraph	wGT3X-BT (Ametris); Actiwatch Spectrum (Philips)	7; 3	Deep learning	MoCA
Liu et al [35]	Commercial	Accelerometer	W180 (Shenzhen Fitfaith)	14	Machine learning	MoCA
Xiao et al [61]	Research	Actigraph	Actigraph GT3XP (Ametris)	7	Statistical	TMT
Hoepel et al [62]	Research	Actigraph	GENEActiv Original (ActivInsights)	4	Statistical	MMSE

<sup>a</sup>MMSE: Mini-Mental State Examination.

<sup>b</sup>Amsterdam IADL: Amsterdam Instrumental Activities of Daily Living.

<sup>c</sup>MoCA: Montreal Cognitive Assessment.

<sup>d</sup>HVLT: Hopkins Verbal Learning Test-Revised.

<sup>e</sup>CDR: Clinical Dementia Rating.

<sup>f</sup>RBANS: Repeatable Battery for the Assessment of Neuropsychological Status.

<sup>g</sup>NINCDS-ADRDA: National Institute of Neurologic, Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders Association.

<sup>h</sup>TMT: Trail-Making Test.

<sup>i</sup>NTB: Neuropsychological Test Battery.

<sup>j</sup>ICD-10: International Statistical Classification of Diseases, Tenth Revision.

<sup>k</sup>RAVLT: Rey Auditory Verbal Learning Test.

<sup>l</sup>PHES: Psychometric Hepatic Encephalopathy Score.

<sup>m</sup>GDS: Global Deterioration Scale.

<sup>n</sup>NIH: National Institutes of Health.

<sup>o</sup>SNSB: Seoul Neuropsychological Screening Battery.

<sup>p</sup>B-SEVLT-Sum: Brief Spanish-English Verbal Learning Test.

**Table 6.** Summary of machine learning used across studies, including model type, main performance metrics, analytical task (classification, regression, or clustering), and the corresponding feature selection or explainable artificial intelligence (XAI) techniques applied. Only the best-performing metric for each model is reported.

Source	Model	Metrics (best results)	Task	Feature selection/XAI
Lentzen et al [22]	LR <sup>a</sup> ; DT <sup>b</sup> ; RF <sup>c</sup> ; XGBoost <sup>d</sup>	AUC <sup>e</sup> =0.73	Binary classification	SHAP <sup>f</sup>
Rykov et al [67]	LR; RF; XGBoost	$R^2=0.690$ ; $\rho=0.700$ ; MAE <sup>g</sup> =0.460	Regression	Correlation-based feature selection
Cho et al [63]	LR; RF; GBM <sup>h</sup> ; SVM <sup>i</sup>	AUC=0.929; accuracy=0.935; precision=0.800; sensitivity=0.956; $F_1$ -score (0.819)	Binary classification	Permutation feature importance
Ghosal et al [38]	GLM <sup>j</sup> ; SOFR <sup>k</sup> ; SOTDR <sup>l</sup> ; SOTDR-L <sup>m</sup>	AUC=0.811; $R^2=0.333$	Binary classification; regression	LASSO <sup>n</sup> /GEL <sup>o</sup> penalties + functional coefficients
Hossain et al [70]	GBM; SVM; LR + LASO; CTGAN <sup>p</sup>	Accuracy=0.948	Multiclass classification; regression	Correlation-based feature selection; wrapper feature selection
Corbi and Burgos [42]	Expectation-maximization clustering	Accuracy=0.910	Unsupervised clustering	Feature relevance via variable testing
Liu et al [35]	GBDT <sup>q</sup> ; XGBoost <sup>r</sup>	Accuracy=0.757; recall=0.952; AUC=0.628	Binary classification	Permutation importance

<sup>a</sup>LR: logistic regression.

<sup>b</sup>DT: decision tree.

<sup>c</sup>RF: random forest.

<sup>d</sup>XGBoost: Extreme Gradient Boosting.

<sup>e</sup>AUC: area under the curve.

<sup>f</sup>SHAP: Shapley additive explanations.

<sup>g</sup>MAE: mean absolute error.

<sup>h</sup>GBM: gradient boosting machine.

<sup>i</sup>SVM: support vector machine.

<sup>j</sup>GLM: generalized linear model.

<sup>k</sup>SOFR: scalar-on-function regression.

<sup>l</sup>SOTDR: scalar on time-by-distribution regression.

<sup>m</sup>SOTDR-L: SOTDR via time varying L moments.

<sup>n</sup>LASSO: least absolute shrinkage and selection operator.

<sup>o</sup>GEL: group exponential LASSO.

<sup>p</sup>CTGAN: conditional tabular generative adversarial network.

<sup>q</sup>GBDT: gradient boosting decision tree.

<sup>r</sup>XGBoost: Extreme Gradient Boosting.

**Table 7.** Summary of deep learning used across studies, including model type, main performance metrics, analytical task (classification, regression, or survival analysis), and the corresponding feature selection or explainable artificial intelligence (XAI) techniques applied. Only the best-performing metric for each model is reported.

Source	Model	Metrics (best results)	Task	Feature selection/XAI
Sun et al [44]	CNN <sup>a</sup> ; ElasticNet; RSF <sup>b</sup>	C-index=0.840; AUC <sup>c</sup> =0.880	Survival analysis (time to Alzheimer disease onset)	Gradient and Gini feature importance; hazard ratio interpretability
Khosroazad et al [69]	Neural network	AUC=0.880; sensitivity=0.870; specificity=0.890	Binary classification	Intrinsic via time-latency
Palmer et al [29]	MS-GAN <sup>d</sup> ; Bayesian LR <sup>e</sup>	Dice=0.730; OR <sup>f</sup> =1.830	Segmentation; regression	Bayesian coefficients (odds ratio and CI)
Lim et al [30]	Neural network + PCA <sup>g</sup>	AUC=0.990	Binary classification	Correlation-based interpretability
Park et al [39]	1D convolutional autoencoder + LR	$R^2=0.979$	Regression	Backward feature elimination
Shi et al [60]	CDPred <sup>h</sup> ; XGBoost <sup>i</sup>	Hip: accuracy=0.84 and AUC=0.86; Wrist: accuracy=0.69 and AUC=0.73	Binary classification	Nonzero predictor relative importance ranking

<sup>a</sup>CNN: convolutional neural network.

<sup>b</sup>RSF: random survival forest.

<sup>c</sup>AUC: area under the curve.

<sup>d</sup>MS-GAN: multispecies generative adversarial network.

<sup>e</sup>LR: logistic regression.

<sup>f</sup>OR: odds ratio.

<sup>g</sup>PCA: principal component analysis.

<sup>h</sup>CDPred: cognitive decline predictor.

<sup>i</sup>XGBoost: Extreme Gradient Boosting.

**Table 8.** Summary of statistical approaches applied across studies, including model type, reporting metrics, covariates or adjustment factors, and main associations between wearable-derived variables and cognitive impairment or dementia indicators. Only the most relevant and statistically significant results for each study are presented. For wearable-derived metrics, arrows indicate the direction of association: an upward arrow (↑) denotes that higher values of the variable are associated with greater cognitive impairment, whereas a downward arrow (↓) indicates the opposite.

Source	Model	Metrics	Score
Espinosa et al [36]	General linear models	CWP <sup>a</sup>	IV <sup>b</sup> (CWP<.001, ↑); IS <sup>c</sup> (CWP<.001, ↓)
Manning et al [37]	Pearson correlation analysis	Pearson CC <sup>d</sup> , <i>P</i> value	Activity counts (CC=-0.829; <i>P</i> =.041, ↓)
Jones et al [23]	Linear regression models	β coefficient, 95% CI	IV (β=.40, <i>q</i> =.022, ↑); sleep onset (β=-.28, 95% CI -0.55 to -0.02, ↑); IS (β=-.27, 95% CI -0.54 to 0.00, ↓)
Satomi et al [24]	Multinomial logistic regression	RR <sup>e</sup> , 95% CI, <i>P</i> value	IV (RR 3.21, 95% CI 1.40-7.34, <i>P</i> =.006, ↑); IS (RR 0.44, 95% CI 0.21-0.93, <i>P</i> =.03, ↓); M10 (RR 0.40, 95% CI 0.18-0.89, <i>P</i> =.02, ↓)
Baril et al [43]	Linear regression models	Standardized β coefficient, <i>P</i> value	Sleep duration (β=.384, <i>P</i> =.001, ↑)
Sauers et al [25]	Linear regression models	Estimate (β), <i>P</i> value	Sleep efficiency (β=-6.026, <i>P</i> <.001, ↓); sleep latency (β=11.302, <i>P</i> <.001, ↑); number of awakenings (β=6.585, <i>P</i> =.001, ↑)
Winer et al [45]	Cox proportional hazards models	HR <sup>g</sup> , 95% CI, <i>P</i> value	IS (HR 1.25, 95% CI 1.050-1.480; <i>P</i> =.007, ↑); amplitude (HR 0.79, 95% CI 0.650-0.960, <i>P</i> =.02, ↓); M10 (HR 0.75, 95% CI 0.610-0.940, <i>P</i> =.01, ↓); MESOR <sup>h</sup> (HR 0.78, 95% CI 0.590-0.998, <i>P</i> =.048, ↓)
Haghighyegh et al [46]	Cox proportional hazard models	HR, 95% CI	Amplitude (HR 1.32, 95% CI 1.17-1.49, ↑); M10 (HR 1.28, 95% CI 1.14-1.44, ↑); L5 (HR 1.15, 95% CI 1.10-1.21, ↑); IV (HR 1.14, 95% CI 1.05-1.24, ↑); rest-activity rhythm (HR 1.23, 95% CI 1.16-1.29, ↑)
Basta et al [47]	ANCOVA	<i>P</i> value	Sleep duration (night TST <sup>j</sup> , <i>P</i> <.001, ↑); TiB <sup>k</sup> (night TiB, <i>P</i> <.001, ↑)
Baril et al [26]	Linear regression models	<i>P</i> value	Sleep duration ( <i>P</i> <.05, ↑); activity counts ( <i>P</i> <.05, ↑); circadian rhythm ( <i>P</i> <.05, ↑)
Milton et al [48]	Multinomial logistic regression	OR <sup>l</sup> , 95% CI	Wake after sleep onset (OR 2.26, 95% CI 1.12-4.55, ↑); sleep efficiency (OR 2.15, 95% CI 1.03-4.47, ↓)
Seo et al [27]	Two-way ANCOVA	<i>P</i> value	Movement/acceleration ( <i>P</i> =.03, ↓)
Huang et al [28]	Unconditional multivariable logistic regression	AOR <sup>m</sup> , 95% CI	MESOR (AOR 1.99, 95% CI 1.04-3.81, ↑)
Skourti et al [49]	Path models (analysis of moment structures)	Standardized β coefficient, <i>P</i> value	Sleep efficiency (direct β=.266, <i>P</i> =.001, ↓); wake after sleep onset (direct β=-.211, <i>P</i> =.001, ↑); TiB (24-hour TiB, indirect β=-.079, <i>P</i> <.001, ↑)
David et al [68]	Spearman rank correlation	Partial correlation coefficient (ρ), <i>P</i> value	Moderate-to-vigorous physical activity (ρ=0.558, <i>P</i> =.02, ↓)
Gao et al [50]	Cox proportional hazards	HR, 95% CI	Amplitude (HR 1.94, 95% CI 1.53-2.46, <i>P</i> <.001, ↑); IV (HR 1.49, 95% CI 1.18-1.88, <i>P</i> <.001, ↑); sleep duration (HR 1.28, 95% CI 1.06-1.55, <i>P</i> =.01, ↑)
Jeon et al [51]	Multivariate linear regressions	β coefficient, <i>P</i> value	Acrophase (β=-.256, <i>P</i> =.004, ↑)
Plotogea et al [64]	Multivariate logistic regression	OR, 95% CI, <i>P</i> value	Sleep efficiency (OR 0.803, 95% CI 0.711-0.904, <i>P</i> =.001, ↓); sleep latency (OR 1.212, 95% CI 1.063-1.383, <i>P</i> =.004, ↑)
Boa Sorte Silva et al [31]	Linear regression models	Unstandardized β (β), <i>P</i> value	Fragmentation index (β=.004, <i>P</i> =.046, ↑)
Palmer et al [40]	Fixel-wise linear regression; Bayesian multiple linear regression	β coefficient, <i>P</i> value	L5 (β=-.29, <i>P</i> <.050, ↑)
Kim et al [32]	Multivariable linear models	Partial rank correlation (ρ), <i>P</i> value	IV (ρ=-0.44, <i>P</i> <.001, ↑); M10 (ρ=0.45, <i>P</i> <.001, ↓); IS (ρ=0.40, <i>P</i> =.009, ↓)
Targa et al [65]	Linear regression models	Effect size, <i>P</i> value	IV (effect size=-0.715, <i>P</i> =.013, ↑)
Wei et al [41]	Descriptive statistics; 1-tailed <i>t</i> test	Mean (SD), <i>P</i> value	Amplitude (0.93, SD 0.59, <i>P</i> =.030, ↓); IS (0.32, SD 0.19, <i>P</i> =.02, ↓); acrophase (44, SD 145, <i>P</i> <.001, ↓)
Cho et al [66]	Generalized linear mixed model	OR, 95% CI, <i>P</i> value	Sleep duration (OR 0.9, 95% CI 0.8-1.0, <i>P</i> <.001, ↓); activity counts (OR 0.02, 95% CI 0.0-0.5, <i>P</i> =.02, ↓)

Source	Model	Metrics	Score
Basta et al [33]	ANCOVA	Mean (SD), <i>P</i> value	Sleep duration (night TST=7.7-hour vs 7.2-hour, <i>P</i> =.011, ↑); sleep duration (24-hour TST=8.5-hour vs 7.8-hour, <i>P</i> =.012, ↑)
Yi Lee et al [52]	Multivariate logistic regression; multinomial logistic regression	AOR, 95% CI	Percent rhythm (AOR 0.26, 95% CI 0.08-0.79, ↓)
Roh et al [34]	Multiple linear regression	Estimate ( $\beta$ ), SE, <i>P</i> value	MESOR ( $\beta$ =1.17, SE=0.37, <i>P</i> <.001, ↓); L5 ( $\beta$ =3.77, SE=1.22, <i>P</i> <.001, ↓)
Lysen et al [53]	Cox proportional hazards	HR, 95% CI	Sleep latency (HR 1.44, 95% CI 1.13-1.83, ↑); TiB (HR 1.40, 95% CI 1.04-1.88, ↑); sleep efficiency (HR 0.72, 95% CI 0.55-0.93, ↓); wake after sleep onset (HR 1.38, 95% CI 1.10-1.74, ↑)
Agudelo et al [54]	Survey linear regression models	$\beta$ coefficient, <i>P</i> value	Sleep latency ( $\beta$ =-.003, <i>P</i> <.001, ↑); sleep duration ( $\beta$ =-.070, <i>P</i> <.05, ↑)
Ning et al [55]	Cox proportional hazards	HR, 95% CI	Moderate-to-vigorous physical activity (HR 0.69, CI 0.54-0.87, <i>P</i> <.001, ↓)
Ning et al [56]	Cox proportional hazards	HR, 95% CI	Moderate-to-vigorous physical activity (HR 0.60, 95% CI 0.40-0.90, ↓)
Zhao et al [57]	Cox proportional hazards	HR, 95% CI	Sleep duration (HR 0.801, 95% CI 0.717-0.893, ↓)
Lu et al [58]	Logistic regression models ANCOVA	OR, 95% CI	Relative amplitude (OR 1.68, 95% CI 1.12-2.50, ↑)
Chan et al [59]	Cox proportional hazards	HR, 95% CI	Bedtime (HR 1.52, 95% CI 1.22-1.85, ↑)
Xiao et al [61]	Cox proportional hazards	HR, 95% CI	MESOR (HR 2.45, 95% CI 1.52-3.94, ↓)
Hoepel et al [62]	Cox proportional hazards	HR, 95% CI	Sedentary behavior (HR 0.36, 95% CI 0.24-0.55, ↓)

<sup>a</sup>CWP: clusterwise *P* value.

<sup>b</sup>IV: intradaily variability.

<sup>c</sup>IS: interdaily stability.

<sup>d</sup>CC: correlation constant.

<sup>e</sup>RR: relative risk.

<sup>f</sup>M10: most active 10-hour.

<sup>g</sup>HR: hazard ratio.

<sup>h</sup>MESOR: midline estimated statistic of rhythm.

<sup>i</sup>L5: least active 5-hour.

<sup>j</sup>TST: total sleep time.

<sup>k</sup>TiB: time in bed.

<sup>l</sup>OR: odds ratio.

<sup>m</sup>AOR: adjusted odds ratio.

## Feasibility of Quantitative Synthesis

The feasibility of quantitative meta-analysis for early detection depends on whether the included studies estimate a common outcome construct or estimand. Quantitative synthesis requires studies to measure the same underlying construct. In the current literature, early detection is not operationalized as a single measurable outcome but is addressed through multiple, nonequivalent analytical objectives and metrics. The evidence base is characterized by the following dimensions:

**Reported statistical metric:** of the 49 included studies, 13 (26.5%) report performance metrics from AI-based approaches, such as accuracy or area under the curve. The remaining 36 studies report heterogeneous statistical effect measures, including hazard ratios for time-to-event analyses (n=16), odds ratios (ORs) for binary risk (n=8), and beta coefficients or correlation measures for continuous associations (n=12). These

metrics reflect distinct inferential targets and are not interchangeable.

**Device and sensor characteristics:** most studies rely on research-grade actigraphy with access to raw accelerometry data (43/49, 87.8%), whereas a smaller subset uses consumer-grade wearables integrating additional sensors and proprietary processing (7/49, 14.3%). Differences in device type, sensor modality, and preprocessing lead to wearable-derived measures that do not map onto a common estimand.

**Study design:** study designs were divided between cross-sectional analyses assessing concurrent associations (21/49, 42.8%) and longitudinal cohort studies estimating prospective risk or disease onset (24/49, 49%). These designs address different research questions and operate over distinct temporal frameworks.

Clinical endpoint: cognitive outcomes included both categorical clinical diagnoses and continuous screening scores. The MMSE was the most frequently used outcome (18/49, 36.7%), followed by *ICD-10 (International Statistical Classification of Diseases, Tenth Revision)*-based diagnoses and the MoCA (each 7/49, 14.3%). These endpoints differ in scale, sensitivity, and clinical interpretation.

Because early detection is not defined as a single measurable construct across studies, no common estimand can be specified for quantitative pooling. A structured narrative synthesis is therefore adopted to examine how analytical approaches, cognitive assessments, and wearable measurement strategies are used to demonstrate early detection capability across the cognitive continuum.

### Wearable Categories

Wearable devices were categorized into 4 groups according to their primary intended use. Research devices were by far the most common, reported in 87.8% (43/49) of studies, and included systems such as ActiGraph, Actiwatch, and accelerometer loggers (eg, Axivity AX3). These devices were purpose-built and validated for assessing activity, sleep, and circadian rhythms in research and clinical contexts. Commercial everyday wearables, such as Fitbit, Apple Watch, and Xiaomi Mi Band, were used in 14.3% (7/49) of studies; these multipurpose consumer devices are affordable and widely accessible but often provide limited access to raw data. Ad hoc prototypes, developed specifically for individual studies, accounted for 8.2% (4/49) of studies. Finally, multimodal research devices were used in 4.1% (2/49) of studies, typically integrating physiological sensing beyond accelerometry, such as heart rate, electrodermal activity, or peripheral arterial tonometry. Counts are nonexclusive because some studies used multiple device types in the same cohort [19].

Across all categories, wear time was typically about 1 week (median, 7 days), with durations ranging from 3-4 days to several months (up to 182.5 days in 1 longitudinal cohort). Most of the studies (29/49, 59.2%) relied on actigraphy devices to capture rest-activity rhythms and sleep, followed by accelerometer devices (12/49, 25.5%) and smartwatches (7/49, 14.3%).

### Cognitive Outcomes

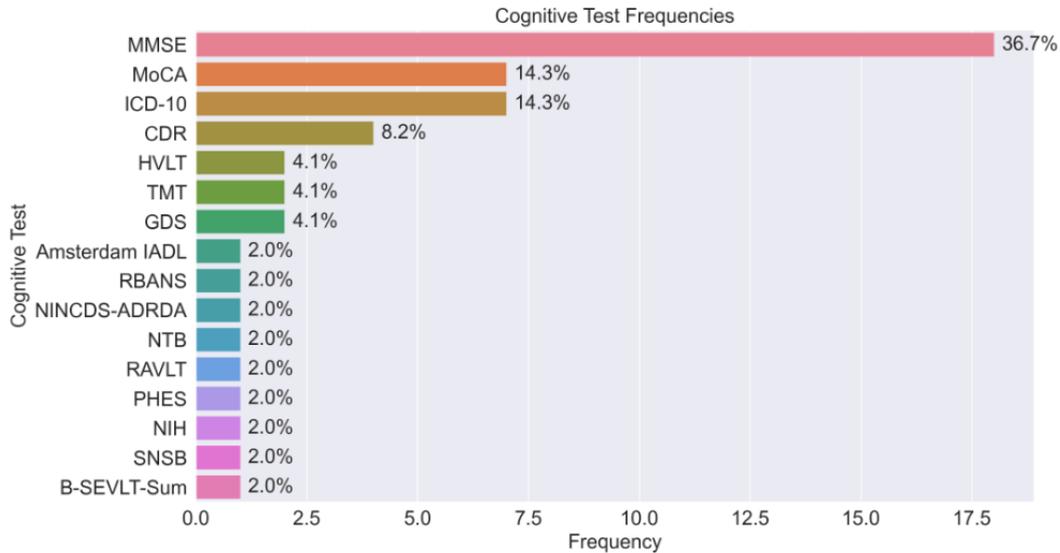
The most frequently used cognitive measures were global screening tests, with their distribution shown in [Figure 3](#). The

MMSE was the most frequently used measure, applied in 36.7% (18/49) of studies. Clinical diagnoses based on *ICD-10* dementia or MCI codes were reported in 14.3% (7/49) of studies. The MoCA appeared in 14.3% (7/49) of studies and Clinical Dementia Rating was used in 8.2% (7/49) of studies. The Hopkins Verbal Learning Test and the Global Deterioration Scale were each used in 4.1% (2/49) of studies. A wide range of other neuropsychological and functional instruments, including the Amsterdam Instrumental Activities of Daily Living scale, Repeatable Battery for the Assessment of Neuropsychological Status, National Institute of Neurologic Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders Association criteria, Trail Making Test, Neuropsychological Test Battery, Rey Auditory Verbal Learning Test, Psychometric Hepatic Encephalopathy Score, National Institutes of Health toolbox, Seoul Neuropsychological Screening Battery, and Brief Spanish-English Verbal Learning Test-Sum, were each applied in 2% (1/49) of studies. The overall distribution of cognitive measures is shown in [Figure 3](#).

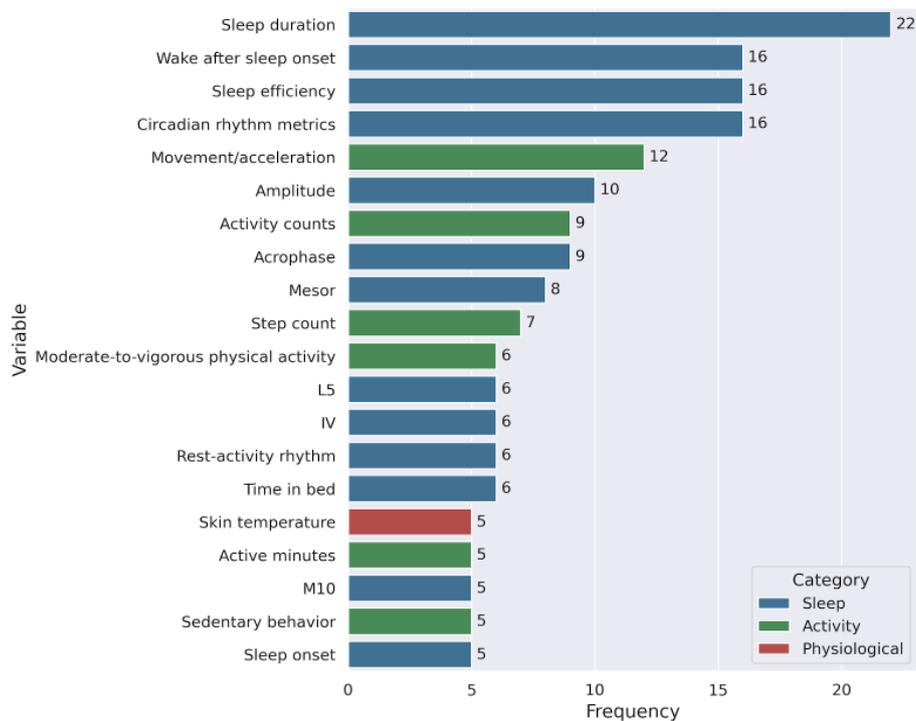
One study applied AI-based approaches to directly predict cognitive test scores [67], while others relied on predefined cut-off points for classification [23,36,41,46,55,57]. Some studies used both direct score estimation and classification approaches [43,54]. A separate line of work focused on predicting survival outcomes, which were not directly comparable to cognitive test scores [43]. Finally, 2 studies addressed related regression tasks by estimating white matter characteristics in the brain from wearable-derived data [44,49].

Across the 49 studies, the most frequently reported wearable-derived variables (complete list is provided in [Multimedia Appendix 6](#)) used to estimate cognitive assessments ([Figure 4](#)) were predominantly sleep-related. Sleep duration was the most frequently reported measure (22/49, 44.9%), followed by wake after sleep onset (16/49, 32.7%), sleep efficiency (16/49, 32.7%), and circadian rhythm metrics (16/49, 32.7%). Activity-related variables were reported less often overall, including movement/acceleration (12/49, 24.5%), activity counts (9/49, 18.4%), step count (7/49, 14.3%), moderate-to-vigorous physical activity (6/49, 12.2%), active minutes (5/49, 10.2%), and sedentary behavior (5/49, 10.2%). Physiological variables were least commonly reported, with skin temperature appearing in a small subset of studies (5/49, 10.2%).

**Figure 3.** Frequency of cognitive assessment instruments across included studies. The bar chart illustrates the distribution of specific cognitive tests and diagnostic criteria used. Frequencies and percentages represent the prevalence of each tool within the total sample of studies. Amsterdam IADL: Amsterdam Instrumental Activities of Daily Living; B-SEVLT-Sum: Brief Spanish-English Verbal Learning Test; CDR: Clinical Dementia Rating; GDS: Global Deterioration Scale; HVL: Hopkins Verbal Learning Test-Revised; ICD-10: International Classification of Diseases, Tenth Revision; MMSE: Mini-Mental State Examination; MoCA: Montreal Cognitive Assessment; NIH: National Institutes of Health; NINCDS-ADRDA: National Institute of Neurologic, Communicative Disorders and Stroke and Alzheimer’s Disease and Related Disorders Association; NTB: Neuropsychological Test Battery; PHES: Psychometric Hepatic Encephalopathy Score; RAVLT: Rey Auditory Verbal Learning Test; RBANS: Repeatable Battery for the Assessment of Neuropsychological Status; SNSB: Seoul Neuropsychological Screening Battery; TMT: Trail-Making Test.



**Figure 4.** Top 20 most frequently used wearable-derived variables across studies. Variables are grouped by domain: sleep (blue), activity (green), and physiological (red). Frequencies represent the number of studies reporting each variable after standardization and filtering for wearable compatibility.



**Wearable Data Analysis Approaches**

Traditional statistical methods were the most frequently applied, reported in 73.5% (36/49) of studies (Table 3). These included group comparisons, correlation analyses, and regression models evaluating associations between wearable-derived activity or sleep metrics and cognitive outcomes. Commonly reported

features were total daily activity counts, sleep duration and efficiency, circadian rhythm indices (intradaily variability, interdaily stability, and relative amplitude), and timing markers such as acrophase or sleep midpoint. The list of wearable-derived features used across these studies, with a description per feature, is included in Multimedia Appendix 6.

Most studies relying on conventional statistical approaches reported significant associations between wearable-derived variables and cognitive outcomes, although the magnitude of reported effects varied with sample size. In smaller studies ( $n < 100$ ), standardized regression coefficients and odds ratios tended to be larger, typically ranging from  $\beta \approx .35-.55$  or OR 2.0-3.1 [24,29,43]. In larger cohort studies ( $n \geq 100$ ), effect estimates were generally lower, around  $\beta \approx .10-.25$  or OR 1.3-1.8 [45,46,49,54].

According to established benchmarks for social and behavioral sciences, standardized coefficients are typically categorized as small ( $\beta \approx .20$ ), medium ( $\beta \approx .50$ ), or large ( $\beta \approx .80$ ) [71]. While the estimates from larger cohorts in this review fall within the small-to-moderate range, they must be interpreted within the specific context of cognitive research. Given that cognitive performance is a multifactorial outcome influenced by a broad array of genetic, environmental, and demographic variables, the observation of stable effect sizes between .10 and .25 is highly meaningful. Such values represent robust contributions to the variance that, despite their conservative magnitude, suggest that wearable-derived metrics are reliable markers of cognitive health at a population level. While most associations reached statistical significance ( $P < .05$ ), only a limited number of studies reported fully adjusted models or external validation [72].

Machine learning methods were applied in 14.3% (7/49) of studies, most often for classification tasks such as distinguishing MCI or dementia from healthy controls. Machine learning models included logistic regression, decision trees, support vector machines, random forests, and gradient-boosting approaches such as Extreme Gradient Boosting (XGBoost). Deep learning approaches were reported in 12.2% (6/49) of studies and were typically applied in larger datasets, where models captured temporal patterns of activity or sleep. Deep learning architectures included neural networks, convolutional neural networks, generative adversarial network-based approaches, and autoencoders. In addition, a small number of studies (2/49, 4.1%) applied clustering techniques to identify subgroups with distinct rest-activity profiles.

Across machine learning and deep learning studies, reported performance metrics varied widely depending on task, cohort size, and study design. Classification performance ranged from moderate discrimination in large, pragmatic settings (eg, area under the receiver operating characteristic curve [AUROC]  $\approx 0.73$ ;  $n = 229$ ; dementia = 160 and healthy controls = 69 [23]) to higher values in smaller or more constrained cohorts (eg, AUROC  $\approx 0.93$ ;  $n = 222$ ; affective symptoms = 126 and not affective symptoms = 96 [63]). Deep learning models reported metrics such as a C-index of approximately 0.84 ( $n = 1077$ ; dementia = 270 and healthy controls = 807) for long-horizon prediction of Alzheimer disease onset [44] and  $R^2$  values approaching 0.98 ( $n = 14,659$ ; dementia = 177 and healthy controls = 14,482) for regression tasks estimating brain structural characteristics from wearable-derived data [40]. Several high-performing models were trained on small or imbalanced datasets, including studies with limited case numbers or highly unequal case-control ratios [28,31,40,70].

In addition, machine learning approaches applied to wearable-derived features achieved high discriminative performance across multiple tasks. Reported results included classification accuracies up to 93.5% [63] ( $n = 222$ ; sleep and nighttime behaviors = 81 and not sleep and nighttime behaviors = 141) for identifying high-risk individuals, accuracies of approximately 94%-95% ( $n = 33$ ; dementia = 5, MCI = 6, and healthy controls = 22) across multiple cognitive impairment levels [70], area-under-the-curve values ranging from approximately 0.73 ( $n = 229$ ; dementia = 160 and healthy controls = 69) [23] to 0.99 ( $n = 18$ ; dementia = 4 and healthy controls = 14) [31] for early or prodromal Alzheimer disease detection, sensitivities and specificities exceeding 85% ( $n = 95$ ; MCI = 50 and healthy controls = 45) [69] for distinguishing MCI from normal cognition, and C-index values up to 0.84 ( $n = 1077$ ; dementia = 207 and healthy controls = 807) [44] for long-horizon Alzheimer disease onset prediction.

Interpretability approaches were reported in a limited subset of machine learning and deep learning studies. Explicit explainability techniques were applied in 46.1% (6/13) of AI-based studies, including Shapley additive explanations values or permutation feature importance in 2 machine learning studies [23,49], functional coefficients in 1 statistical-machine learning hybrid approach [31], and Bayesian coefficient estimation in 1 deep learning framework [32]. One additional study reported feature relevance through statistical testing in a clustering-based approach [39]. The remaining AI-based studies relied on feature-selection heuristics or dimensionality reduction methods, such as correlation-based selection, wrapper methods, or principal component analysis, without reporting interpretable model outputs [33,35,67,70].

### Prevention-Oriented Findings

The included studies were divided into 2 categories according to the type of contribution reported. A total of 22.4% (11/49) of studies explicitly provided quantitative results relevant to the early detection or prevention of cognitive decline (direct evidence), defined as studies demonstrating or validating predictive or preventive applications. The remaining 38 (77.6%) studies contributed indirect evidence, identifying associations between wearable-derived features and cognitive outcomes that may inform, but do not yet constitute, preventive or predictive applications.

Among studies providing direct evidence ( $n = 11$ ), longitudinal or interventional designs were commonly used. Specifically, 45.5% (5/11) of studies used longitudinal follow-up or randomized designs, while 54.5% (6/11) of studies relied on cross-sectional analyses. Direct-evidence studies included randomized intervention trials reporting outcomes such as slower progression to dementia, improved gait speed, or increased adherence to preventive programs, as well as longitudinal cohort studies showing that disrupted rest-activity rhythms, sleep fragmentation, and physical activity patterns predicted incident dementia or MCI over follow-up periods extending up to 8-15 years. Several studies further demonstrated that wearable-derived motor activity features could forecast clinical Alzheimer disease onset, with reported C-index values ranging from approximately 0.80 to 0.84.

In contrast, studies contributing indirect evidence ( $n=38$ ) more frequently adopted cross-sectional designs. Of these, 52.6% (20/38) of studies were cross-sectional, whereas 47.4% (18/38) of studies used longitudinal designs. These studies primarily focused on identifying associations between wearable-derived sleep, circadian rhythm, activity, or physiological measures and cognitive outcomes, including global cognitive scores, neuropsychological performance, biomarkers, or clinical diagnoses, rather than on prediction or intervention.

Across all included studies, longitudinal designs accounted for 44.9% (22/49) of studies, while 55.1% (27/49) of studies used cross-sectional designs. This distribution reflects differing study aims, with direct evidence studies more often incorporating longitudinal follow-up or intervention components, and indirect evidence studies emphasizing associative analyses.

Overall, wearable technologies contributed to prevention-oriented research both directly, by demonstrating predictive and early detection capabilities in longitudinal and interventional settings, and indirectly, by identifying behavioral and physiological markers associated with future cognitive decline. This dual contribution reflects the role of wearables as both measurement tools for preventive interventions and sources of early behavioral risk markers in dementia research.

## Discussion

### Principal Findings

This systematic review synthesized 49 studies published since 2020 on wearable devices for the early detection and prevention of cognitive impairment and dementia. The findings are presented in 4 dimensions corresponding to the subobjectives described earlier.

### Wearable Devices and Measurement Contexts

Research-grade actigraphy and accelerometry devices were used in most of the included studies (43/49, 87.8%), reflecting both a strong methodological emphasis on validated access to raw accelerometry data and the relative maturity of actigraphy-based approaches in this field [73,74]. Evidence from these studies was derived predominantly from observational designs, including cross-sectional and longitudinal cohorts, and was frequently assessed as having low risk of bias, particularly among cohort studies with large samples [48,53,60]. Across actigraphy-based studies, sample sizes were heterogeneous but predominantly large, with nearly four-fifths enrolling 100 participants or more, and large samples were strongly associated with low risk of bias. While research-grade devices enable robust investigation of behavioral markers associated with cognitive impairment, their limited scalability and accessibility may constrain broader clinical and public health deployment [75].

In contrast, consumer-grade wearable devices were comparatively underrepresented despite their widespread adoption and potential for large-scale, longitudinal monitoring. Evidence from consumer-wearable studies was derived predominantly from cross-sectional designs and exhibited greater heterogeneity in sample size and methodological quality, with smaller studies more frequently assessed as having

moderate risk of bias, reflecting restricted access to raw data and reliance on proprietary signal processing [41,68,76]. However, consumer-wearable evidence also included a small number of large cross-sectional studies assessed as having low risk of bias, indicating that methodological rigor is driven more by study scale and design than by device class alone [22,34,35]. Taken together, these findings suggest that different wearable device classes capture complementary aspects of behavior and physiology rather than interchangeable measurements, supporting the potential value of multimodal integration or purpose-driven device selection aligned with specific research or clinical objectives [76].

### Wearable-Derived Features and Cognitive Outcomes

Across included studies, cognition was frequently assessed using brief global screening instruments or standardized clinical diagnostic classifications, reflecting a pragmatic choice aligned with clinical practice and the feasibility requirements of observational research. Evidence derived from commonly used cognitive measures was drawn primarily from cross-sectional and cohort designs and was frequently assessed as having low risk of bias, particularly in larger cohort studies [46,50,55,56]. While global cognitive measures facilitate standardized comparisons across heterogeneous cohorts, their limited sensitivity to subtle or domain-specific changes constrains interpretation of associations with wearable-derived features, especially in preclinical populations [77].

Wearable devices were not used to measure cognition directly but to capture continuous behavioral and physiological signals that were statistically associated with cognitive outcomes. Sleep-related and circadian rhythm measures were examined most frequently and were reported across both cross-sectional and longitudinal studies, predominantly in cohort designs with large samples, and were most often assessed as having low risk of bias when derived from validated actigraphy devices and predefined metrics [33,47,59,62]. In contrast, physiological signals beyond accelerometry were examined in only a small subset of studies, which used heterogeneous, primarily noncohort designs, often involved small to medium sample sizes, and showed higher proportions of moderate-to-high-risk-of-bias assessments [35,41] limiting the strength of inferences that can currently be drawn from these measures and highlighting priorities for future methodological development.

### Analytical Approaches and Methodological Considerations

Conventional statistical methods were the predominant analytical approach, used in 73.5% (36/49) of studies, reflecting their interpretability and familiarity in clinical and epidemiological research. These methods were applied mainly in cross-sectional and longitudinal observational designs and were frequently assessed as having low risk of bias, particularly in large cohort studies using multivariable or survival models [45,46,50,55-62]. Across these studies, associations between wearable-derived features and cognitive outcomes were generally modest and sensitive to covariate adjustment, especially in large and heterogeneous samples, underscoring

the importance of adequate sample size, rigorous model specification, and replication for reliable inference.

Machine learning and deep learning approaches were applied in a smaller subset of studies (13/49, 26.5%) and were used primarily for classification or prediction tasks. These studies were frequently exploratory, relied on small or imbalanced samples, and were more often assessed as having moderate to high risk of bias, largely due to limited external validation and optimistic performance estimates [67,69,70]. Across AI-based studies, higher predictive performance was more commonly reported in cross-sectional or nonlongitudinal designs using constrained datasets, whereas applications in larger or longitudinal cohorts generally yielded more modest performance estimates but were more often assessed as having low or moderate risk of bias, reflecting improved robustness and generalizability [35,44]. Overall, this pattern suggests a trade-off between performance and methodological robustness, in which performance gains observed in small or exploratory samples may reflect overfitting or cohort-specific structure rather than generalizable predictive signal.

### Implications for Early Detection and Prevention

Evidence supporting early detection or prevention was derived from a limited subset of studies providing direct prevention evidence, which more often used cohort-based or predictive modeling designs, enrolled larger samples, and were predominantly assessed as having low risk of bias. Representative examples include large longitudinal or predictive studies demonstrating that wearable-derived behavioral measures, particularly sleep-wake organization and circadian regularity, can precede clinical cognitive impairment by several years [39,46,56]. These findings support the potential role of wearables in early risk stratification rather than post hoc characterization.

In contrast, most of the included studies contributed indirect prevention evidence and relied primarily on cross-sectional or association-focused designs. Although these studies were frequently assessed as having low to moderate risk of bias and consistently reported associations between wearable-derived markers and cognitive status, their typically limited temporal resolution and lack of predictive validation constrained their ability to establish preventive relevance [24,27,33,40]. As a result, their contribution to prevention remains inferential rather than actionable.

Overall, the current evidence supports wearable devices as tools for monitoring and early risk identification rather than as stand-alone preventive interventions [78]. Across both direct and indirect evidence, disrupted sleep-wake patterns and circadian irregularity consistently emerged as markers of elevated cognitive risk [33,39,55]. Translating these associations into effective prevention strategies will require confirmation in larger, longitudinal, and externally validated studies [78] that integrate wearable-based risk stratification with clinical decision-making and behavioral intervention frameworks.

### Limitations

Several limitations should be considered when interpreting these findings. The included studies were heterogeneous in design,

target populations, device types, and outcome definitions. Devices ranged from consumer-grade wearables to research actigraphy and multimodal physiological instruments, and the metrics derived from them were inconsistently defined. Cognitive outcomes also varied: global screening tools such as the MMSE and MoCA were most frequent, but other studies used different neuropsychological tests, functional measures, or clinical diagnoses. This heterogeneity precluded formal meta-analysis and necessitated a structured narrative synthesis. We did not formally assess the certainty of evidence (eg, using Grading of Recommendations Assessment, Development, and Evaluation), given the methodological heterogeneity of included studies and the predominance of observational designs.

The evidence base is also constrained by small sample sizes and short monitoring periods. In total, 8.2% (4/49) of studies enrolled 30 participants or fewer, while an additional 24.5% (12/49) included between 30 and 100 participants, indicating that more than one-third of the evidence relies on small cohorts. Furthermore, 67.3% (33/49) of studies monitored participants for only 1 week or less, limiting the ability to capture long-term variability in activity, sleep, or circadian rhythms. Only 18.4% (9/49) of studies systematically assessed real-world feasibility, including adherence, comfort, or long-term usability. Moreover, external validation was performed in just 6.1% (3/49) of studies, restricting the generalizability of reported associations or predictive models beyond the original cohorts. Thus, a key limitation of the current evidence is the risk of overinterpretation. Most studies were cross-sectional or small cohorts, and only few have undergone external validation [29,39,63]. In several studies, high accuracy or AUROC values were reported in the context of small or imbalanced datasets, conditions that increase the risk of overfitting and optimism bias. As a result, these performance estimates should be interpreted cautiously in the absence of external validation or evaluation in independent cohorts. While models often report high accuracy, these findings may not generalize. Preventive claims remain speculative until tested in large prospective cohorts or intervention trials. In addition, research-grade devices typically provide access to raw, high-resolution data and validated measurement pipelines, whereas consumer devices often rely on proprietary algorithms and offer limited transparency, which may affect comparability across studies.

The risk of bias was present to varying degrees. Although most studies were rated as low risk (34/49, 69.4%), nearly one-third (13/49, 26.5%) were classified as medium risk, and 4.1% (2/49) as high risk. The high-risk studies exemplify common problems such as selective sampling, inadequate reference standards, and small or restricted populations, all of which likely led to overestimation of diagnostic performance and limited the generalizability of their findings. Publication bias is also possible, as studies reporting positive or novel results are more likely to be published. Together, these issues highlight that the field remains fragmented, with methodological and reporting inconsistencies that restrict the strength and reproducibility of current evidence.

Limitations of the review process itself must also be acknowledged. The restriction to English-language publications and peer-reviewed journals may have introduced language or

publication bias, potentially excluding relevant gray literature or local studies. Additionally, while we searched for 4 major interdisciplinary databases, the exclusion of specialized clinical indices (eg, PsycINFO or Embase) could have resulted in the omission of some relevant records.

### Implications for Future Research and Practice

This review highlights several priorities for future research and practice. A key challenge is bridging the gap between research-grade actigraphy and consumer wearables. While actigraphy remains the most validated tool for assessing sleep, activity, and circadian rhythms, limited use of consumer devices constrains scalability [79]. Evidence suggests that some multisensor consumer wearables can achieve accuracy comparable to research-grade actigraphs under controlled conditions [80,81], but direct comparisons across diverse populations and real-world settings remain needed. Where sufficient accuracy and reliability are demonstrated, consumer wearables could support large-scale, longitudinal monitoring for research and preventive care [82].

Another priority is the expanded use of advanced analytics and digital phenotyping. Longitudinal wearable data enable modeling of dynamic behavioral and physiological changes, supporting early and personalized prevention. However, machine learning and deep learning methods remain underused, particularly sequence-based models such as recurrent neural networks, temporal convolutional models, and transformer-based approaches (eg, Informer and TimesFM) that are well suited for time-series health data [83]. In addition, large language model-based frameworks (eg, HealthLLM) offer opportunities to integrate wearable-derived time series with clinical and contextual information [84]. Progress in this area will depend on access to larger, well-annotated datasets and multicenter collaboration, as well as the incorporation of explainable AI methods to ensure interpretability and clinical trust [85].

Finally, more prevention-oriented research is needed. Only 2 randomized controlled trials tested wearable-assisted interventions, and a small subset of observational studies examined prevention-relevant behaviors. Yet prevention is central to dementia care given the absence of disease-modifying treatments. Wearables could play a dual role, both by detecting early signs of decline and by monitoring modifiable lifestyle factors such as sleep, circadian rhythms, and physical activity [86,87]. Positioning wearables within preventive frameworks could therefore be among the most clinically meaningful directions for future research.

From a translational perspective, ethical and regulatory considerations will be critical for the clinical adoption of AI-based dementia risk prediction. Issues such as data privacy, transparency of algorithms, potential psychological impact of early risk labeling, and compliance with regulatory frameworks must be carefully addressed. At present, research-grade wearable devices and analytically transparent models appear more suitable for clinical research settings, whereas many consumer-grade devices and advanced AI approaches remain primarily experimental. Further validation, standardization, and regulatory oversight will be necessary before these methods can be routinely integrated into clinical practice.

Overall, this review advances the field by framing the transition from descriptive statistics to a predictive digital phenotyping framework. Distinct from prior reviews limited to established dementia, it synthesizes evidence specifically for the preclinical window, distinguishing findings with direct predictive utility from those offering only indirect correlational insights. It contributes a critical assessment of methodological maturity, identifying the heavy reliance on research-grade devices and the lack of external validation as primary barriers to implementation. Ultimately, these findings suggest that shifting to continuous, passive monitoring offers a scalable method to detect subtle behavioral deviations, creating opportunities for earlier, personalized risk reduction strategies [88].

---

### Acknowledgments

The authors thank the Vicomtech Foundation for supporting this research. The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing, reformatting, and quality assessment. The GAI tool used was: ChatGPT-5.1 (OpenAI). Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by: “collective responsibility.”

---

### Funding

This research was funded by the Vicomtech Foundation. The funder had no role in the study design, data collection, analysis, interpretation, or writing of the manuscript.

---

### Data Availability

No new primary data were generated in this systematic review. Data extracted from the included studies and the curated list of screened articles are available from the corresponding author upon reasonable request. The Python code used to support automated screening and study management is not publicly available, as it was developed for internal use; however, it can be shared upon reasonable request for purposes of methodological transparency.

---

## Authors' Contributions

Conceptualization was carried out by AC. Methodology was developed by AC and MA, and software was prepared by AC. Validation was conducted by CM and AA. Data curation was performed by AC and MA. The original draft of the manuscript was written by AC and MA, and review and editing were undertaken by AC, MA, CM, and AA. Supervision was provided by CM and AA, while project administration was managed by CM. Funding acquisition was secured by CM. All authors have read and agreed to the published version of the manuscript.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

PRISMA 2020 checklist.

[\[DOCX File, 40 KB-Multimedia Appendix 1\]](#)

## Multimedia Appendix 2

PRISMA-S checklist.

[\[DOCX File, 17 KB-Multimedia Appendix 2\]](#)

## Multimedia Appendix 3

Exact search queries used across all databases to identify records included in this review.

[\[PDF File \(Adobe PDF File\), 53 KB-Multimedia Appendix 3\]](#)

## Multimedia Appendix 4

Extended study characteristics and principal findings of all the studies included in this systematic review.

[\[XLSX File \(Microsoft Excel File\), 16 KB-Multimedia Appendix 4\]](#)

## Multimedia Appendix 5

Risk of bias assessments of included studies using validated tools appropriate to study design.

[\[PDF File \(Adobe PDF File\), 276 KB-Multimedia Appendix 5\]](#)

## Multimedia Appendix 6

List and definitions of sleep, circadian rhythm, physical activity, and physiological variables extracted from wearable devices across included studies.

[\[PDF File \(Adobe PDF File\), 111 KB-Multimedia Appendix 6\]](#)

## References

1. Dementia. World Health Organization. 2025. URL: <https://www.who.int> [accessed 2025-08-30]
2. Chowdhary N, Barbui C, Anstey KJ, Kivipelto M, Barbera M, Peters R. Reducing the risk of cognitive decline and dementia: WHO recommendations. *Front Neurol*. 2021;12:765584. [[FREE Full text](#)] [doi: [10.3389/fneur.2021.765584](https://doi.org/10.3389/fneur.2021.765584)] [Medline: [35082745](https://pubmed.ncbi.nlm.nih.gov/35082745/)]
3. Risk Reduction of Cognitive Decline and Dementia: WHO guidelines. Geneva. World Health Organization; 2019:96.
4. Chong TWH, Macpherson H, Schaumberg MA, Brown BM, Naismith SL, Steiner GZ. Dementia prevention: the time to act is now. *Med J Aust*. 2021;214(7):302-304. [doi: [10.5694/mja2.50972](https://doi.org/10.5694/mja2.50972)] [Medline: [33715184](https://pubmed.ncbi.nlm.nih.gov/33715184/)]
5. Jones A, Ali MU, Kenny M, Mayhew A, Mokashi V, He H. Potentially modifiable risk factors for dementia and mild cognitive impairment: an umbrella review and meta-analysis. *Dement Geriatr Cogn Disord*. 2024;53(2):91-106. [[FREE Full text](#)] [doi: [10.1159/000536643](https://doi.org/10.1159/000536643)] [Medline: [38346414](https://pubmed.ncbi.nlm.nih.gov/38346414/)]
6. Hardcastle C, Taylor B, Price C. Global Deterioration Scale. In: *Encyclopedia of Gerontology and Population Aging*. Cham. Springer; 2022:2198-2201.
7. Baron-Shahaf D, Shahaf G. Markers of too little effort or too much alertness during neuropsychological assessment: demonstration with perioperative changes. *Brain Behav*. 2024;14(8):e3649. [[FREE Full text](#)] [doi: [10.1002/brb3.3649](https://doi.org/10.1002/brb3.3649)] [Medline: [39169455](https://pubmed.ncbi.nlm.nih.gov/39169455/)]
8. Holding BC, Ingre M, Petrovic P, Sundelin T, Axelsson J. Quantifying cognitive impairment after sleep deprivation at different times of day: a proof of concept using ultra-short smartphone-based tests. *Front Behav Neurosci*. 2021;15:666146. [doi: [10.3389/fnbeh.2021.666146](https://doi.org/10.3389/fnbeh.2021.666146)] [Medline: [33927603](https://pubmed.ncbi.nlm.nih.gov/33927603/)]

9. Salis F, Costagiu D, Mandas A. Mini-mental state examination: optimal cut-off levels for mild and severe cognitive impairment. *Geriatrics (Basel)*. 2023;8(1):12. [FREE Full text] [doi: [10.3390/geriatrics8010012](https://doi.org/10.3390/geriatrics8010012)] [Medline: [36648917](https://pubmed.ncbi.nlm.nih.gov/36648917/)]
10. Arevalo-Rodriguez I, Smailagic N, Roqué-Figuls M, Ciapponi A, Sanchez-Perez E, Giannakou A. Mini-mental state examination (MMSE) for the early detection of dementia in people with mild cognitive impairment (MCI). *Cochrane Database Syst Rev*. 2021;7(7):CD010783. [doi: [10.1002/14651858.CD010783.pub3](https://doi.org/10.1002/14651858.CD010783.pub3)] [Medline: [34313331](https://pubmed.ncbi.nlm.nih.gov/34313331/)]
11. Chithiramohan T, Santhosh S, Threlfall G, Hull L, Mukaetova-Ladinska EB, Subramaniam H. Culture-fair cognitive screening tools for assessment of cognitive impairment: a systematic review. *J Alzheimers Dis Rep*. 2024;8(1):289-306. [doi: [10.3233/adr-230194](https://doi.org/10.3233/adr-230194)]
12. Sakal C, Li T, Li J, Li X. Predicting poor performance on cognitive tests among older adults using wearable device data and machine learning: a feasibility study. *NPJ Aging*. 2024;10(1):56. [FREE Full text] [doi: [10.1038/s41514-024-00177-x](https://doi.org/10.1038/s41514-024-00177-x)] [Medline: [39587119](https://pubmed.ncbi.nlm.nih.gov/39587119/)]
13. Soumma SB, Mamun A, Ghasemzadeh H. AI-powered wearable sensors for health monitoring and clinical decision making. *Curr Opin Biomed Eng*. 2025;36:100628. [doi: [10.1016/j.cobme.2025.100628](https://doi.org/10.1016/j.cobme.2025.100628)]
14. Keshet A, Reicher L, Bar N, Segal E. Wearable and digital devices to monitor and treat metabolic diseases. *Nature metabolism*. 2023;5(4):563-571. [doi: [10.1038/s42255-023-00778-y](https://doi.org/10.1038/s42255-023-00778-y)]
15. Zaman Q, others. The role of artificial intelligence in early disease detection: transforming diagnostics and treatment. *Multidiscip J Healthc (MJH)*. 2024:1-54. [FREE Full text]
16. Cote AC, Phelps RJ, Kabiri NS, Bhangu JS, Thomas. Evaluation of wearable technology in dementia: a systematic review and meta-analysis. *Front Med (Lausanne)*. 2021;7:501104. [FREE Full text] [doi: [10.3389/fmed.2020.501104](https://doi.org/10.3389/fmed.2020.501104)] [Medline: [33505979](https://pubmed.ncbi.nlm.nih.gov/33505979/)]
17. Weizman Y, Tirosch O, Beh J, Fuss FK, Pedell S. Gait assessment using wearable sensor-based devices in people living with dementia: a systematic review. *Int J Environ Res Public Health*. 2021;18(23):12735. [doi: [10.3390/ijerph182312735](https://doi.org/10.3390/ijerph182312735)] [Medline: [34886459](https://pubmed.ncbi.nlm.nih.gov/34886459/)]
18. Iso-Markku P, Aaltonen S, Kujala UM, Halme H, Phipps D, Knittle K. Physical activity and cognitive decline among older adults: a systematic review and meta-analysis. *JAMA Netw Open*. 2024;7(2):e2354285. [doi: [10.1001/jamanetworkopen.2023.54285](https://doi.org/10.1001/jamanetworkopen.2023.54285)] [Medline: [38300618](https://pubmed.ncbi.nlm.nih.gov/38300618/)]
19. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. [FREE Full text] [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
20. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Syst Rev*. 2021;10(1):39. [FREE Full text] [doi: [10.1186/s13643-020-01542-z](https://doi.org/10.1186/s13643-020-01542-z)] [Medline: [33499930](https://pubmed.ncbi.nlm.nih.gov/33499930/)]
21. Ma LL, Wang YY, Yang ZH, Huang D, Weng H, Zeng X. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? *Mil Med Res*. 2020;7(1):7. [doi: [10.1186/s40779-020-00238-8](https://doi.org/10.1186/s40779-020-00238-8)] [Medline: [32111253](https://pubmed.ncbi.nlm.nih.gov/32111253/)]
22. Lentzen M, Vairavan S, Muurling M, Alepopoulos V, Atreya A, Boada M. RADAR-AD: assessment of multiple remote monitoring technologies for early detection of Alzheimer's disease. *Alzheimers Res Ther*. 2025;17(1):29. [doi: [10.1186/s13195-025-01675-0](https://doi.org/10.1186/s13195-025-01675-0)] [Medline: [39865315](https://pubmed.ncbi.nlm.nih.gov/39865315/)]
23. Jones CD, Wasilko R, Zhang G, Stone KL, Gujral S, Rodakowski J. Detecting sleep/wake rhythm disruption related to cognition in older adults with and without mild cognitive impairment using the myRhythmWatch platform: feasibility and correlation study. *JMIR Aging*. 2025;8:e67294-e67294. [doi: [10.2196/67294](https://doi.org/10.2196/67294)] [Medline: [40194310](https://pubmed.ncbi.nlm.nih.gov/40194310/)]
24. Satomi E, Apolinário D, Magaldi RM, Busse AL, Vieira Gomes GC, Ribeiro E. Beyond sleep: rest and activity rhythm as a marker of preclinical and mild dementia in older adults with less education. *Neurobiol Sleep Circadian Rhythms*. 2025;18:100110. [FREE Full text] [doi: [10.1016/j.nbscr.2024.100110](https://doi.org/10.1016/j.nbscr.2024.100110)] [Medline: [39834590](https://pubmed.ncbi.nlm.nih.gov/39834590/)]
25. Sauers SC, Toedebusch CD, Richardson R, Spira AP, Morris JC, Holtzman DM. Midpoint of sleep is associated with sleep quality in older adults with and without symptomatic Alzheimer's disease. *Sleep Adv*. 2024;5(1):zpa023. [doi: [10.1093/sleepadvances/zpa023](https://doi.org/10.1093/sleepadvances/zpa023)] [Medline: [38711547](https://pubmed.ncbi.nlm.nih.gov/38711547/)]
26. Baril AA, Picard C, Labonté A, Sanchez E, Duclos C, Mohammediyani B. Day-to-day sleep variability with Alzheimer's biomarkers in at-risk elderly. *Alzheimers Dement Diagn Assess Dis Monit*. 2024;16(1):e12521. [FREE Full text] [doi: [10.1002/dad2.12521](https://doi.org/10.1002/dad2.12521)] [Medline: [38371359](https://pubmed.ncbi.nlm.nih.gov/38371359/)]
27. Seo K, Takayanagi N, Sudo M, Yamashiro Y, Chiba I, Makino K. Association between daily gait speed patterns and cognitive impairment in community-dwelling older adults. *Sci Rep*. 2023;13(1):2783. [doi: [10.1038/s41598-023-29805-4](https://doi.org/10.1038/s41598-023-29805-4)] [Medline: [36797381](https://pubmed.ncbi.nlm.nih.gov/36797381/)]
28. Huang B, Liao G, Lee PMY, Chan CK, Tai L, Tsang CYJ. Association of circadian rhythm with mild cognitive impairment among male pneumoconiosis workers in Hong Kong: a cross-sectional study. *Sci Rep*. 2023;13(1):1650. [doi: [10.1038/s41598-023-28832-5](https://doi.org/10.1038/s41598-023-28832-5)] [Medline: [36717700](https://pubmed.ncbi.nlm.nih.gov/36717700/)]
29. Palmer JR, Wang C, Kong D, Cespedes M, Pye J, Hickie IB. Rest-activity rhythms and tract specific white matter lesions in older adults at risk for cognitive decline. *Mol Psychiatry*. 2022;27(8):3410-3416. [FREE Full text] [doi: [10.1038/s41380-022-01641-4](https://doi.org/10.1038/s41380-022-01641-4)] [Medline: [35764707](https://pubmed.ncbi.nlm.nih.gov/35764707/)]

30. Lim J. A smart healthcare-based system for classification of dementia using deep learning. *Digit Health*. 2022;8:20552076221131667. [FREE Full text] [doi: [10.1177/20552076221131667](https://doi.org/10.1177/20552076221131667)] [Medline: [36312848](https://pubmed.ncbi.nlm.nih.gov/36312848/)]
31. Boa Sorte Silva NC, Falck RS, Chan PCY, Tai D, Backhouse D, Stein R. The association of sleep and cortical thickness in mild cognitive impairment. *Experimental Gerontology*. 2022;167:111923. [doi: [10.1016/j.exger.2022.111923](https://doi.org/10.1016/j.exger.2022.111923)]
32. Kim M, Liotta EM, Maas MB, Braun RI, Garcia-Canga B, Ganger DR. Rest-activity rhythm disturbance in liver cirrhosis and association with cognitive impairment. *Sleep*. 2021;44(6):zsaa288. [doi: [10.1093/sleep/zsaa288](https://doi.org/10.1093/sleep/zsaa288)] [Medline: [33367862](https://pubmed.ncbi.nlm.nih.gov/33367862/)]
33. Basta M, Zaganas I, Simos P, Koutentaki E, Dimovasilis C, Mathioudakis L. Apolipoprotein E  $\epsilon 4$  allele is associated with long sleep duration among elderly with cognitive impairment. *J Alzheimers Dis*. 2021;79(2):763-771. [doi: [10.3233/jad-200958](https://doi.org/10.3233/jad-200958)]
34. Roh HW, Choi J, Kim NR, Choe YS, Choi JW, Cho SM. Associations of rest-activity patterns with amyloid burden, medial temporal lobe atrophy, and cognitive impairment. *EBioMedicine*. 2020;58:102881. [doi: [10.1016/j.ebiom.2020.102881](https://doi.org/10.1016/j.ebiom.2020.102881)] [Medline: [32736306](https://pubmed.ncbi.nlm.nih.gov/32736306/)]
35. Liu Z, Zhang L, Wu J, Zheng Z, Gao J, Lin Y. Machine learning-based classification of circadian rhythm characteristics for mild cognitive impairment in the elderly. *Front Public Health*. 2022;10:1036886. [doi: [10.3389/fpubh.2022.1036886](https://doi.org/10.3389/fpubh.2022.1036886)] [Medline: [36388285](https://pubmed.ncbi.nlm.nih.gov/36388285/)]
36. Espinosa N, Hoyos CM, McKinnon AC, Almgren H, Duffy SL, Naismith SL. Rest-activity rhythm fragmentation and synchronization are linked with reduced cortical thickness in older adults "at risk" for dementia. *Sleep*. 2025;48(5):zsaf017. [doi: [10.1093/sleep/zsaf017](https://doi.org/10.1093/sleep/zsaf017)] [Medline: [40052961](https://pubmed.ncbi.nlm.nih.gov/40052961/)]
37. Manning CA, Youngkin A, Quigg M. Sleepiness, sleep time, and depression of caregivers are linked with sleep and behaviors of their paired partners with dementia. *J Alzheimers Dis Rep*. 2025;9:25424823241300981. [FREE Full text] [doi: [10.1177/25424823241300981](https://doi.org/10.1177/25424823241300981)] [Medline: [40034516](https://pubmed.ncbi.nlm.nih.gov/40034516/)]
38. Ghosal R, Varma VR, Volfson D, Urbanek J, Hausdorff JM, Watts A. Scalar on time-by-distribution regression and its application for modelling associations between daily-living physical activity and cognitive functions in Alzheimer's Disease. *Sci Rep*. 2022;12(1):11558. [doi: [10.1038/s41598-022-15528-5](https://doi.org/10.1038/s41598-022-15528-5)] [Medline: [35798763](https://pubmed.ncbi.nlm.nih.gov/35798763/)]
39. Park B, Choi BJ, Lee H, Jang JH, Roh HW, Kim EY. Modeling brain volume using deep learning-based physical activity features in patients with dementia. *Front Neuroinform*. 2022;16:795171. [FREE Full text] [doi: [10.3389/fninf.2022.795171](https://doi.org/10.3389/fninf.2022.795171)] [Medline: [35356447](https://pubmed.ncbi.nlm.nih.gov/35356447/)]
40. Palmer JR, Duffy SL, Meares S, Pye J, Calamante F, Cespedes M. Rest-activity functioning is related to white matter microarchitecture and modifiable risk factors in older adults at-risk for dementia. *Sleep*. 2021;44(7):zsab007. [doi: [10.1093/sleep/zsab007](https://doi.org/10.1093/sleep/zsab007)] [Medline: [33428761](https://pubmed.ncbi.nlm.nih.gov/33428761/)]
41. Wei J, Boger J. Sleep detection for younger adults, healthy older adults, and older adults living with dementia using wrist temperature and actigraphy: prototype testing and case study analysis. *JMIR Mhealth Uhealth*. 2021;9(6):e26462. [FREE Full text] [doi: [10.2196/26462](https://doi.org/10.2196/26462)] [Medline: [34061038](https://pubmed.ncbi.nlm.nih.gov/34061038/)]
42. Corbi A, Burgos D. Connection between sleeping patterns and cognitive deterioration in women with alzheimer's disease. *Sleep Breath*. 2022;26(1):361-371. [doi: [10.1007/s11325-021-02327-x](https://doi.org/10.1007/s11325-021-02327-x)] [Medline: [33792886](https://pubmed.ncbi.nlm.nih.gov/33792886/)]
43. Baril A, Picard C, Labonté A, Sanchez E, Duclos C, Mohammediyani B. Longer sleep duration and neuroinflammation in at-risk elderly with a parental history of Alzheimer's disease. *Sleep*. 2024;47(6):zsae081. [FREE Full text] [doi: [10.1093/sleep/zsae081](https://doi.org/10.1093/sleep/zsae081)] [Medline: [38526098](https://pubmed.ncbi.nlm.nih.gov/38526098/)]
44. Sun H, Li P, Gao L, Yang J, Yu L, Buchman AS. Altered motor activity patterns within 10-minute timescale predict incident clinical Alzheimer's disease. *J Alzheimers Dis*. 2024;98(1):209-220. [FREE Full text] [doi: [10.3233/JAD-230928](https://doi.org/10.3233/JAD-230928)] [Medline: [38393904](https://pubmed.ncbi.nlm.nih.gov/38393904/)]
45. Winer JR, Lok R, Weed L, He Z, Poston KL, Mormino EC. Impaired 24-h activity patterns are associated with an increased risk of Alzheimer's disease, Parkinson's disease, and cognitive decline. *Alzheimers Res Ther*. 2024;16(1):35. [FREE Full text] [doi: [10.1186/s13195-024-01411-0](https://doi.org/10.1186/s13195-024-01411-0)] [Medline: [38355598](https://pubmed.ncbi.nlm.nih.gov/38355598/)]
46. Haghayegh S, Gao C, Sugg E, Zheng X, Yang HW, Saxena R. Association of rest-activity rhythm and risk of developing dementia or mild cognitive impairment in the middle-aged and older population: prospective cohort study. *JMIR Public Health Surveill*. 2024;10:e55211. [FREE Full text] [doi: [10.2196/55211](https://doi.org/10.2196/55211)] [Medline: [38713911](https://pubmed.ncbi.nlm.nih.gov/38713911/)]
47. Basta M, Bouloukaki I, Skourti E, Zampetakis A, Alexopoulou C, Ganiaris A. Long objective sleep duration is a marker of cognitive impairment in older adults: findings from the Cretan Aging Cohort. *J Alzheimers Dis Rep*. 2024;8(1):927-934. [FREE Full text] [doi: [10.3233/ADR-230203](https://doi.org/10.3233/ADR-230203)] [Medline: [38910938](https://pubmed.ncbi.nlm.nih.gov/38910938/)]
48. Milton S, Cavaillès C, Ancoli-Israel S, Stone KL, Yaffe K, Leng Y. Five-year changes in 24-hour sleep-wake activity and dementia risk in oldest old women. *Neurology*. 2025;104(8):e213403. [doi: [10.1212/WNL.0000000000213403](https://doi.org/10.1212/WNL.0000000000213403)] [Medline: [40106755](https://pubmed.ncbi.nlm.nih.gov/40106755/)]
49. Skourti E, Simos P, Zampetakis A, Koutentaki E, Zaganas I, Alexopoulou C. Long-term associations between objective sleep quality and quantity and verbal memory performance in normal cognition and mild cognitive impairment. *Front Neurosci*. 2023;17:1265016. [doi: [10.3389/fnins.2023.1265016](https://doi.org/10.3389/fnins.2023.1265016)] [Medline: [37928739](https://pubmed.ncbi.nlm.nih.gov/37928739/)]
50. Gao L, Li P, Gaykova N, Zheng X, Gao C, Lane JM. Circadian rest-activity rhythms, delirium risk, and progression to dementia. *Ann Neurol*. 2023;93(6):1145-1157. [FREE Full text] [doi: [10.1002/ana.26617](https://doi.org/10.1002/ana.26617)] [Medline: [36808743](https://pubmed.ncbi.nlm.nih.gov/36808743/)]

51. Jeon SY, Byun MS, Yi D, Jung G, Lee J, Kim YK. Circadian rest-activity rhythm and longitudinal brain changes underlying late-life cognitive decline. *Psychiatry Clin Neurosci*. 2023;77(4):205-212. [FREE Full text] [doi: [10.1111/pcn.13521](https://doi.org/10.1111/pcn.13521)] [Medline: [36527292](https://pubmed.ncbi.nlm.nih.gov/36527292/)]
52. Yi Lee PM, Ling Kwok BH, Ting Ma JY, Tse LA. A population-based prospective study on rest-activity rhythm and mild cognitive impairment among Hong Kong healthy community-dwelling older adults. *Neurobiol Sleep Circadian Rhythms*. 2021;10:100065. [FREE Full text] [doi: [10.1016/j.nbscr.2021.100065](https://doi.org/10.1016/j.nbscr.2021.100065)] [Medline: [33997474](https://pubmed.ncbi.nlm.nih.gov/33997474/)]
53. Lysen TS, Luik AI, Ikram MK, Tiemeier H, Ikram MA. Actigraphy-estimated sleep and 24-hour activity rhythms and the risk of dementia. *Alzheimers Dement*. 2020;16(9):1259-1267. [FREE Full text] [doi: [10.1002/alz.12122](https://doi.org/10.1002/alz.12122)] [Medline: [32558256](https://pubmed.ncbi.nlm.nih.gov/32558256/)]
54. Agudelo C, Tarraf W, Wu B, Wallace DM, Patel SR, Redline S. Actigraphic sleep patterns and cognitive decline in the hispanic community health study/study of latinos. *Alzheimers Dement*. 2021;17(6):959-968. [doi: [10.1002/alz.12250](https://doi.org/10.1002/alz.12250)] [Medline: [33350583](https://pubmed.ncbi.nlm.nih.gov/33350583/)]
55. Ning Y, Chen M, Yang H, Jia J. Accelerometer-derived "weekend warrior" physical activity, sedentary behavior, and risk of dementia. *Alzheimers Res Ther*. 2025;17(1):67. [FREE Full text] [doi: [10.1186/s13195-024-01657-8](https://doi.org/10.1186/s13195-024-01657-8)] [Medline: [40121491](https://pubmed.ncbi.nlm.nih.gov/40121491/)]
56. Ning Y, Chen M, Yang H, Jia J. Accelerometer-measured physical activity timing with incident dementia. *Alzheimers Dement*. 2025;21(2):e14452. [doi: [10.1002/alz.14452](https://doi.org/10.1002/alz.14452)] [Medline: [39868637](https://pubmed.ncbi.nlm.nih.gov/39868637/)]
57. Zhao B, Zhou S, Chang J, Li A, Geng C, Wei T. Association between weekend recovery sleep and risk of incident dementia: a prospective cohort study in the UK Biobank. *J Neurol*. 2025;272(9):612. [doi: [10.1007/s00415-025-13363-y](https://doi.org/10.1007/s00415-025-13363-y)] [Medline: [40911069](https://pubmed.ncbi.nlm.nih.gov/40911069/)]
58. Lu Z, Leung JCS, Feng H, Zhang J, Wing YK, Kwok TCY. Circadian rest-activity rhythms and cognitive decline and impairment in older Chinese adults: a multicohort study with prospective follow-up. *Arch Gerontol Geriatr*. 2023;116:105215. [doi: [10.1016/j.archger.2023.105215](https://doi.org/10.1016/j.archger.2023.105215)] [Medline: [39491073](https://pubmed.ncbi.nlm.nih.gov/39491073/)]
59. Chan LLY, Espinoza Cerda MT, Brodie MA, Lord SR, Taylor ME. Daily-life walking speed, running duration and bedtime from wrist-worn sensors predict incident dementia: a watch walk - UK biobank study. *Int Psychogeriatr*. 2025;37(3):100031. [FREE Full text] [doi: [10.1016/j.inpsyc.2024.100031](https://doi.org/10.1016/j.inpsyc.2024.100031)] [Medline: [39827009](https://pubmed.ncbi.nlm.nih.gov/39827009/)]
60. Shi C, Babiker N, Urbanek JK, Grossman RL, Huisingh-Scheetz M, Rzhetsky A. Free-living wrist and hip accelerometry forecast cognitive decline among older adults without dementia over 1- or 5-years in two distinct observational cohorts. *NPJ Aging*. 2022;8(1):7. [doi: [10.1038/s41514-022-00087-w](https://doi.org/10.1038/s41514-022-00087-w)] [Medline: [35927250](https://pubmed.ncbi.nlm.nih.gov/35927250/)]
61. Xiao Q, Shadyab AH, Rapp SR, Stone KL, Yaffe K, Sampson JN. Rest-activity rhythms and cognitive impairment and dementia in older women: results from the women's health initiative. *J Am Geriatr Soc*. 2022;70(10):2925-2937. [FREE Full text] [doi: [10.1111/jgs.17926](https://doi.org/10.1111/jgs.17926)] [Medline: [35708069](https://pubmed.ncbi.nlm.nih.gov/35708069/)]
62. Hoepel SJW, Ikram MK, Ikram MA, Voortman T, Luik AI. Substitutions of sleep, physical activity, and sedentary behaviour and the risk of dementia and stroke. *J Sleep Res*. 2026;35(1):e70166. [doi: [10.1111/jsr.70166](https://doi.org/10.1111/jsr.70166)] [Medline: [40793962](https://pubmed.ncbi.nlm.nih.gov/40793962/)]
63. Cho E, Kim S, Heo S, Shin J, Hwang S, Kwon E. Machine learning-based predictive models for the occurrence of behavioral and psychological symptoms of dementia: model development and validation. *Sci Rep*. 2023;13(1):8073. [FREE Full text] [doi: [10.1038/s41598-023-35194-5](https://doi.org/10.1038/s41598-023-35194-5)] [Medline: [37202454](https://pubmed.ncbi.nlm.nih.gov/37202454/)]
64. Plotogea OM, Diaconu CC, Gheorghe G, Stan-Ilie M, Badea M, Prelipcean CC. The prevalence and association of cognitive impairment with sleep disturbances in patients with chronic liver disease. *Brain Sci*. 2022;12(4):444. [doi: [10.3390/brainsci12040444](https://doi.org/10.3390/brainsci12040444)] [Medline: [35447976](https://pubmed.ncbi.nlm.nih.gov/35447976/)]
65. Targa ADS, Benítez ID, Dakterzada F, Fontenele-Araujo J, Minguez O, Zetterberg H. The circadian rest-activity pattern predicts cognitive decline among mild-moderate alzheimer's disease patients. *Alzheimers Res Ther*. 2021;13(1):161. [FREE Full text] [doi: [10.1186/s13195-021-00903-7](https://doi.org/10.1186/s13195-021-00903-7)] [Medline: [34563258](https://pubmed.ncbi.nlm.nih.gov/34563258/)]
66. Cho E, Kim S, Hwang S, Kwon E, Heo S, Lee JH, et al. Factors associated with behavioral and psychological symptoms of dementia: prospective observational study using actigraphy. *J Med Internet Res*. 2021;23(10):e29001. [doi: [10.2196/29001](https://doi.org/10.2196/29001)] [Medline: [34714244](https://pubmed.ncbi.nlm.nih.gov/34714244/)]
67. Rykov YG, Patterson MD, Gangwar BA, Jabar SB, Leonardo J, Ng KP. Predicting cognitive scores from wearable-based digital physiological features using machine learning: data from a clinical trial in mild cognitive impairment. *BMC Med*. 2024;22(1):36. [FREE Full text] [doi: [10.1186/s12916-024-03252-y](https://doi.org/10.1186/s12916-024-03252-y)] [Medline: [38273340](https://pubmed.ncbi.nlm.nih.gov/38273340/)]
68. David S, Hohenfeld C, Haberl L, Pahl J, Costa AS, Kilders A. Physical activity monitoring in Alzheimer's disease during sport interventions: a multi-methodological perspective. *Front Neurol*. 2023;14:1195694. [FREE Full text] [doi: [10.3389/fneur.2023.1195694](https://doi.org/10.3389/fneur.2023.1195694)] [Medline: [37808485](https://pubmed.ncbi.nlm.nih.gov/37808485/)]
69. Khosroozad S, Gilbert CF, Aronis JB, Daigle KM, Esfahani M, Almaghasilah A. Sleep movements and respiratory coupling as a biobehavioral metric for early Alzheimer's disease in independently dwelling adults. *BMC Geriatr*. 2023;23(1):252. [FREE Full text] [doi: [10.1186/s12877-023-03983-2](https://doi.org/10.1186/s12877-023-03983-2)] [Medline: [37106470](https://pubmed.ncbi.nlm.nih.gov/37106470/)]
70. Hossain M, Uddin M, Thapa K, Al Zubaer MA, Islam M, Lee J. Detecting cognitive impairment status using keystroke patterns and physical activity data among the older adults: a machine learning approach. *J Healthc Eng*. 2021;2021:1302989. [doi: [10.1155/2021/1302989](https://doi.org/10.1155/2021/1302989)] [Medline: [34966518](https://pubmed.ncbi.nlm.nih.gov/34966518/)]
71. Kallogjeri D, Spitznagel EL, Piccirillo JF. Importance of defining and interpreting a clinically meaningful difference in clinical research. *JAMA Otolaryngol Head Neck Surg*. 2020;146(2):101-102. [doi: [10.1001/jamaoto.2019.3744](https://doi.org/10.1001/jamaoto.2019.3744)] [Medline: [31804662](https://pubmed.ncbi.nlm.nih.gov/31804662/)]

72. Rykov YG, Patterson MD, Gangwar BA, Jabar SB, Leonardo J, Ng KP. Predicting cognitive scores from wearable-based digital physiological features using machine learning: data from a clinical trial in mild cognitive impairment. *BMC Med*. 2024;22(1):36. [FREE Full text] [doi: [10.1186/s12916-024-03252-y](https://doi.org/10.1186/s12916-024-03252-y)] [Medline: [38273340](https://pubmed.ncbi.nlm.nih.gov/38273340/)]
73. Berger A, Wielgus KK, Young-McCaughan S, Fischer P, Farr L, Lee KA. Methodological challenges when using actigraphy in research. *J Pain Symptom Manage*. 2008;36(2):191-199. [FREE Full text] [doi: [10.1016/j.jpainsymman.2007.10.008](https://doi.org/10.1016/j.jpainsymman.2007.10.008)] [Medline: [18400460](https://pubmed.ncbi.nlm.nih.gov/18400460/)]
74. Patterson MR, Nunes AAS, Gerstel D, Pilkar R, Guthrie T, Neishabouri A. 40 years of actigraphy in sleep medicine and current state of the art algorithms. *NPJ Digit Med*. 2023;6(1):51. [FREE Full text] [doi: [10.1038/s41746-023-00802-1](https://doi.org/10.1038/s41746-023-00802-1)] [Medline: [36964203](https://pubmed.ncbi.nlm.nih.gov/36964203/)]
75. Morris AC, Seker A, Telesia L, Wickersham A, Ching BC, Roy R. Adherence to actigraphic devices in elementary school-aged children: systematic review and meta-analysis. *J Med Internet Res*. 2025;27:e79718. [FREE Full text] [doi: [10.2196/79718](https://doi.org/10.2196/79718)] [Medline: [41183377](https://pubmed.ncbi.nlm.nih.gov/41183377/)]
76. Dobson R, Stowell M, Warren J, Tane T, Ni L, Gu Y. Use of consumer wearables in health research: issues and considerations. *J Med Internet Res*. 2023;25:e52444. [FREE Full text] [doi: [10.2196/52444](https://doi.org/10.2196/52444)] [Medline: [37988147](https://pubmed.ncbi.nlm.nih.gov/37988147/)]
77. D'Ignazio G, Carlucci L, Sergi MR, Palumbo R, Dattilo L, Terrei M. Is the MMSE enough for MCI? A narrative review of the usefulness of the MMSE. *Front Psychol*. 2025;16:1727738. [FREE Full text] [doi: [10.3389/fpsyg.2025.1727738](https://doi.org/10.3389/fpsyg.2025.1727738)] [Medline: [41356018](https://pubmed.ncbi.nlm.nih.gov/41356018/)]
78. Gu B, Kim HS, Kim H, Yoo J. Advancements in wearable sensor technologies for health monitoring in terms of clinical applications, rehabilitation, and disease risk assessment: systematic review. *JMIR Mhealth Uhealth*. 2026;14:e76084. [doi: [10.2196/76084](https://doi.org/10.2196/76084)] [Medline: [41511829](https://pubmed.ncbi.nlm.nih.gov/41511829/)]
79. Kainec KA, Caccavaro J, Barnes M, Hoff C, Berlin A, Spencer RMC. Evaluating accuracy in five commercial sleep-tracking devices compared to research-grade actigraphy and polysomnography. *Sensors (Basel)*. 2024;24(2):635. [FREE Full text] [doi: [10.3390/s24020635](https://doi.org/10.3390/s24020635)] [Medline: [38276327](https://pubmed.ncbi.nlm.nih.gov/38276327/)]
80. Smith MT, McCrae CS, Cheung J, Martin JL, Harrod CG, Heald JL, et al. Use of actigraphy for the evaluation of sleep disorders and circadian rhythm sleep-wake disorders: an American academy of sleep medicine systematic review, meta-analysis, and GRADE assessment. *J Clin Sleep Med*. 2018;14(7):1209-1230. [doi: [10.5664/jcsm.7228](https://doi.org/10.5664/jcsm.7228)] [Medline: [29991438](https://pubmed.ncbi.nlm.nih.gov/29991438/)]
81. Kahawage P, Jumabhoy R, Hamill K, de Zambotti M, Drummond SPA. Validity, potential clinical utility, and comparison of consumer and research-grade activity trackers in Insomnia Disorder I: in-lab validation against polysomnography. *J Sleep Res*. 2020;29(1):e12931. [doi: [10.1111/jsr.12931](https://doi.org/10.1111/jsr.12931)] [Medline: [31626361](https://pubmed.ncbi.nlm.nih.gov/31626361/)]
82. Huhn S, Axt M, Gunga HC, Maggioni MA, Munga S, Obor D. The impact of wearable technologies in health research: scoping review. *JMIR Mhealth Uhealth*. 2022;10(1):e34384. [FREE Full text] [doi: [10.2196/34384](https://doi.org/10.2196/34384)] [Medline: [35076409](https://pubmed.ncbi.nlm.nih.gov/35076409/)]
83. Ahmed S, Nielsen IE, Tripathi A, Siddiqui S, Ramachandran RP, Rasool G. Transformers in time-series analysis: a tutorial. *Circuits Syst Signal Process*. 2023;42(12):7433-7466. [doi: [10.1007/s00034-023-02454-8](https://doi.org/10.1007/s00034-023-02454-8)]
84. Kim Y, Xu X, McDuff D, Breazeal C, Park HW. Healthllm: large language models for health prediction via wearable sensor data. *ArXiv*. Preprint posted online on January 12, 2024. [doi: [10.48550/arXiv.2401.06866](https://doi.org/10.48550/arXiv.2401.06866)]
85. Saraswat D, Bhattacharya P, Verma A, Prasad VK, Tanwar S, Sharma G. Explainable AI for healthcare 5.0: opportunities and challenges. *IEEE Access*. 2022;10:84486-84517. [doi: [10.1109/access.2022.3197671](https://doi.org/10.1109/access.2022.3197671)]
86. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet*. 2020;396(10248):413-446. [FREE Full text] [doi: [10.1016/S0140-6736\(20\)30367-6](https://doi.org/10.1016/S0140-6736(20)30367-6)] [Medline: [32738937](https://pubmed.ncbi.nlm.nih.gov/32738937/)]
87. Kivipelto M, Mangialasche F, Ngandu T. Lifestyle interventions to prevent cognitive impairment, dementia and Alzheimer disease. *Nat Rev Neurol*. 2018;14(11):653-666. [doi: [10.1038/s41582-018-0070-3](https://doi.org/10.1038/s41582-018-0070-3)] [Medline: [30291317](https://pubmed.ncbi.nlm.nih.gov/30291317/)]
88. Bufano P, Laurino M, Said S, Tognetti A, Menicucci D. Digital phenotyping for monitoring mental disorders: systematic review. *J Med Internet Res*. 2023;25:e46778. [FREE Full text] [doi: [10.2196/46778](https://doi.org/10.2196/46778)] [Medline: [38090800](https://pubmed.ncbi.nlm.nih.gov/38090800/)]

## Abbreviations

**AI:** artificial intelligence

**ICD-10:** International Statistical Classification of Diseases, Tenth Revision

**MCI:** mild cognitive impairment

**MeSH:** Medical Subject Headings

**MMSE:** Mini-Mental State Examination

**MoCA:** Montreal Cognitive Assessment

**OR:** odds ratio

**PRISMA:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses

**PRISMA-S:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses literature search extension

**XGBoost:** Extreme Gradient Boosting

*Edited by S Brini; submitted 22.Oct.2025; peer-reviewed by B Howell, S Ajayi, S Ye; comments to author 27.Nov.2025; revised version received 23.Jan.2026; accepted 26.Jan.2026; published 23.Feb.2026*

*Please cite as:*

*Cejudo A, Arrojo M, Martín C, Almeida A*

*AI and Wearables for Early Detection of Cognitive Impairment and Dementia: Systematic Review*

*J Med Internet Res 2026;28:e86262*

*URL: <https://www.jmir.org/2026/1/e86262>*

*doi: [10.2196/86262](https://doi.org/10.2196/86262)*

*PMID:*

©Ander Cejudo, Markel Arrojo, Cristina Martín, Aitor Almeida. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 23.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.